ANNALS OF MATHEMATICS
Vol. 42, No. 4, October, 1941

A PROPERTY OF INTEGRAL FUNCTIONS OF ORDER LESS THAN TWO WITH REAL ROOTS

BY K. S. K. IYENGAR

(Received March 29, 1940)

1. Introduction

P. Erdös and T. Grünwald in their paper "On polynomials with real roots," in *Annals of Mathematics*, Vol. 40, No. 3, July 1939, p. 537, have proved the following: Let $f(x)$ be a real polynomial with real roots having consecutive single roots at $x = a_1$, $x = a_2$, and let P be the point of intersection of the tangents to the curve $Y = f(x)$ at $(a_1, 0)$, $(a_2, 0)$, and let $M(a_1a_2)$ be maximum of $f(x)$ in (a_1a_2). Then

$$(1.1) \qquad \tfrac{2}{3}M(a_1a_2)\,|\,a_2 - a_1\,| \geqq \left|\int_{a_1}^{a_2} f(x)\,dx\right| \geqq \tfrac{1}{3}\,|\,a_2 - a_1\,|\cdot|\,Y(P)\,|$$

where $Y(P)$ is the ordinate of the point P.

It is the object of this paper to give two very comprehensive theorems, and give the best possible inequality (3.3) of the type (1.1), from which it will be obvious that the inequality (1.1) is the best possible only in cases of symmetry.

2. Statement of the theorems

THEOREM 1: *Let $a_1 < a_2$, $m_1 > 0 > -m_2$, and let $\phi(x)$ be an integral function of order less than two, satisfying the following conditions.*

(1) *the roots of $\phi(x) = 0$ are all real*

(2) *$\phi(a_1) = \phi(a_2) = 0$ and $\phi(x) \neq 0$ in $a_1 < x < a_2$*

(3) *$\phi'(a_1) = m_1$ and $\phi'(a_2) = -m_2$.*

Then all the curves $Y = \phi(x)$ lie above the curve

$$(2.1) \qquad Y = -A\cdot e^{B\cdot x}(x - a_1)(x - a_2) = m(x).$$

$$\text{Where } A = (m_1m_2)^{1/2}/a_2 - a_1\cdot e^{-B/2(a_1+a_2)} \text{ and } B = \frac{\log m_2/m_1}{a_2 - a_1}$$

and you can find polynomials with real roots satisfying conditions (2) and (3) as near as you like to the minimum curve (2.1).

THEOREM 2: *Let $a_1 < a_2$ and let $\phi(x)$ be an integral function of order < 2 such that,*

(1) *all roots of $\phi(x) = 0$ are real.*

(2) *$\phi(a_1) = \phi(a_2) = 0$ and $\phi(x) > 0$ in $a_1 < x < a_2$ and $\phi'(a_1) \neq 0$, $\phi'(a_2) \neq 0$*

(3) *$\phi'(x_0) = 0$, $\quad a_1 < x_0 < a_2$*

(4) *$\phi(x_0) = M_0$.*

Then all the curves $Y = \phi(x)$ *lie below the curve*

$$Y = \bar{A}(x - a_1)(a_2 - x)e^{\bar{B}x} = M(x) \tag{2.2}$$

$$\text{where } \bar{B} = \frac{2x_0 - a_1 - a_2}{(x_0 - a_1)(a_2 - x_0)} \quad \text{and} \quad \bar{A} = \frac{M_0 e^{-\bar{B}x_0}}{(x_0 - a_1)(a_2 - x_0)}$$

and you can find polynomials with real roots satisfying (2), (3) *and* (4) *as near as you like to the maximum curve* (2.2).

THEOREM 3_1: *The curve* $Y = Ae^{Bx}(x - a_1)(a_2 - x)$ *of theorem* 1 *as in* 2.1 *lies above the parabola which touches the lines* $Y = m_1(x - a_1)$ *and* $Y = -m_2(x - a_2)$ *at* $(a_1, 0)$ *and* $(a_2, 0)$ *and coincides with the minimum curve in case* $m_1 = m_2$.

THEOREM 3_2: *The curve* $Y = \bar{A}e^{\bar{B}x}(x - a_1)(a_2 - x)$ *of theorem* 2 *as in* 2.2 *lies below the parabola passing through* $(a_1, 0)$, $(a_2, 0)$ *and touching the line* $Y = M_0$ *at* x_0 *and if* $x_0 = \dfrac{a_1 + a_2}{2}$ *then the parabola will coincide with the maximum curve of* 2.2.

From theorems 1 and 3_1 we get

$$\int_{a_1}^{a_2} \phi(x)\,dx \geqq \int_{a_1}^{a_2} m(x)\,dx$$

$$= (a_2 - a_1)^2 \cdot (m_1 m_2)^{1/2} \cdot \left\{\frac{1}{6} + \frac{1}{2}\int_0^{B_1} \frac{u(\sinh u - u)}{B_1^3}\,du\right\} \tag{3.1}$$

$$\left[\text{where } B_1 = \frac{B(a_2 - a_1)}{2},\ B \text{ being defined in 2.1}\right]$$

$$\geqq \text{area of the parabola of contact} = \tfrac{1}{3}\,Y(P)(a_2 - a_1)^*$$

where $Y(P)$ is the Y-coordinate of the point of intersection of the tangents at $(a_1, 0)$, $(a_2, 0)$ to $Y = \phi(x)$.

From Theorems 2 and 3_2 we get

$$\int_{a_1}^{a_2} \phi(x)\,dx \leqq \int_{a_1}^{a_2} M(x)\,dx = \frac{e^{-2u_0^2/1-u_0^2}}{1 - u_0^2} \cdot 2M_0 \cdot (a_2 - a_1) \int_0^{B_2} \frac{u \cdot \sinh u}{B_2^3}\,du \tag{3.2}$$

$$\left[\text{where } u_0 = x_0 - \frac{(a_1 + a_2)}{2} \Big/ \frac{a_2 - a_1}{2} \quad \text{and} \quad B_2 = 2u_0/1 - u_0^2\right]$$

$$\leqq \text{area of the parabola in } (a_1 a_2) \text{ of theorem } 3_2 = \frac{2M_0}{3}(a_2 - a_1)^1$$

---

[1] Note: The two elementary propositions, namely:

(i) if $Y = P(x)$ be a parabola touching the lines $Y = m_1(x - a_1)$ $Y = -m_2(x - a_2)$ at a_1, a_2 respectively

$$\int_{a_1}^{a_2} Y\,dx = \tfrac{1}{3}Y(P)(a_2 - a_1) = \frac{(a_2 - a_1)}{3} \cdot \left(\frac{m_1 m_2}{m_1 + m_2}\right)$$

so that combining (3.1) and (3.2) we get

$$\begin{aligned}\frac{2M_0}{3}(a_2-a_1) &\geqq \frac{e^{-2u_0^3/1-u_0^2}}{1-u_0^2}\cdot 2M_0(a_2-a_1)\cdot\int_0^{B_2}\frac{u\cdot\sinh u}{B_2^3}\,du \\ &\geqq \int_{a_1}^{a_2}\phi(x)\,dx \\ &\geqq (a_2-a_1)^2(m_1m_2)^{1/2}\left\{\frac{1}{6}+\frac{1}{2}\cdot\int_0^{B_1}\frac{u(\sinh u-u)}{B_1^3}\,du\right\} \\ &\geqq \tfrac{1}{3}Y(P)(a_2-a_1)\end{aligned} \tag{3.3}$$

where $\phi(x)$ is any integral function of order < 2 satisfying the following conditions.

(1) $\phi(x) = 0$ are all real.
(2) $\phi(a_1) = \phi(a_2) = 0 \quad \phi(x) > 0 \quad$ in $a_1 < x < a_2$
(3) $\phi'(a_1) = m_1 \quad \phi'(a_2) = -m_2$
(4) $\phi'(x_0) = 0 \quad \phi(x_0) = M_0 \quad a_1 < x_0 < a_2$.

It will be obvious in the course of the proofs of theorems 3_1 and 3_2 that equality in the first and last inequality signs in (3.3) can occur only in the case of symmetry, namely

$$m_1 = m_2 \quad \text{and} \quad x_0 = \frac{a_1+a_2}{2}$$

so that the Erdös-Grünwald inequality is the best possible only in case of symmetry.

NOTE: We may generalize theorem 1 to the case where conditions (2) and (3) are generalized as follows:

(2) $\phi(a_1) = 0$ is a zero of pth degree $\quad \phi(a_2) = 0$ of qth degree
(3) $\phi^p(a_1) = m_1 \quad \phi^q(a_2) = m_2$.

Then the minimum curve will be of the type

$Y = A(x-a_1)^p(a_2-x)^qe^{Bx}$ where A and B will be given in terms of m_1 and m_2.

A similar generalization for theorem 2 is also possible.

3. Proof of theorems

PROOF OF THEOREM 1. Let $F(x) = \dfrac{\phi(x)}{m(x)}$ where $m(x)$ is the function defined in (2.1); then

$$F(a_1) = F(a_2) = 1 \tag{4.1}$$

(ii) If $Y = P(x)$ be a parobola touching the line $Y = M_0$ at x_0 in (a_1a_2) and passing through $(a_1, 0)$, $(a_2, 0)$ then

$$\int_{a_1}^{a_2} y\,dx = \tfrac{2}{3}\cdot M_0(a_2-a_1),$$

are assumed here.

and since $\phi(x)$ is an integral function of order < 2

$$\phi(x) = e^{\alpha x+\beta}(x - a_1)(a_2 - x)\left\{\prod_{n=3}^{\infty}\left(1 - \frac{x}{a_n}\right)e^{\frac{x}{a_n}}\right\} \tag{4.2}$$

all the a_n's for $n \geqq 3$ lying outside (a_1a_2). Hence

$$\frac{d^2}{dx^2}\log F(x) = -\sum_{n=3}^{\infty}\frac{1}{(x - a_n)^2} \tag{4.3}$$

so that $\log F(x)$ is concave (downwards) in the interval (a_1a_2) and hence the arc of $\log F(x)$ will lie above the chord in (a_1a_2) or

$$\log F(x) \geqq 0 \quad \text{or} \quad \phi(x) \geqq m(x) \qquad \text{in } (a_1a_2), \tag{4.4}$$

equality occurring only when $\phi(x)$ coincides with $m(x)$.

To prove the latter half of theorem 1, without any loss of generalization we take $a_1 = -1$, $a_2 = +1$. Let $m(x)$ be the corresponding minimum curve as defined in 2.1.

Let $P(x)$ be the polynomial $P(x) = (1 - x^2)\left(1 - \frac{x}{a}\right)^n$ a and n to be chosen presently, $\{|a| > 1\}$; n being given by $a = \dfrac{\left(\frac{m_2}{m_1}\right)^{\frac{1}{n}} + 1}{\left(\frac{m_2}{m_1}\right)^{\frac{1}{n}} - 1}$ so that $P'(-1) = m_1$, $P'(1) = -m_2$.

Let $m_2 > m_1$ (the argument being the same also when $m_2 < m_1$). Let

$$\theta(x) = \log\frac{P(x)}{m(x)},$$

then $$\theta(\pm 1) = 0 \quad \text{and} \quad \theta''(x) = -\frac{n}{(x - a)^2}.$$

If $-1 < x < 1$

$$|\theta(x)| \leqq \frac{(1 - x^2)}{2}\cdot[\text{Max of } |\theta''(x)| \text{ in } (-1, 1)].$$

Now since $m_2 > m_1$, $(a > 1)$ and max. of $|\theta''|$ in $(-1, 1)$ will be at $x = 1$

$$|\theta''(1)| = \frac{n}{(a - 1)^2} = \frac{n}{4}\cdot\left\{\left(\frac{m_2}{m_1}\right)^{\frac{1}{n}} - 1\right\}^2$$

and for large n the last term will be $O\left(\frac{1}{n}\right)$, hence $P(x) = m(x)e^{o\left(\frac{1}{n}\right)} = m(x)(1 + \epsilon_n)$, thus proving theorem 1 completely.

Proof of Theorem 2: Let $F(x) = \log\dfrac{\phi(x)}{M(x)}$ where $M(x)$ is the maximum

function as defined in (2.2). Then

$$F(x_0) = F'(x_0) = 0,$$

and since ϕ is an integral function of order < 2,

$$\phi(x) = e^{\alpha x+\beta}(x - a_1)(a_2 - x)\cdot\left\{\prod_3^\infty \left(1 - \frac{x}{a_n}\right) e^{\frac{x}{a_n}}\right\}$$

a_n for $n \geqq 3$ lying outside $(a_1 a_2)$.

Since $$\frac{d^2}{dx^2} F(x) = -\sum_3^\infty \frac{1}{(x - a_n)^2} \leqq 0, \qquad \text{in } (a_1 a_2)$$

and $F'(x_0) = 0$ at x_0 in $(a_1 < x_0 < a_2)$ we see that $F(x) \leqq F(x_0) = 0$ in $(a_1 \leqq x \leqq a_2)$, hence $\phi(x) \leqq M(x)$.

The latter part of theorem 2, that we can find polynomials satisfying given condition as near as you like to $M(x)$ can be proved in the same manner as the latter part of theorem 1.

Proof of Theorem 3_1: Without loss of generality we may take $a_1 = -1$ $a_2 = 1$; given $m_2 \neq m_1$, the parabola of contact will be given by

$$Y = \sqrt{\alpha^2 + \beta^2 + 2\alpha\cdot\beta\cdot x} - \alpha - \beta x = P(x)$$

where $$\alpha = \frac{2m_1 m_2(m_1 + m_2)}{(m_1 - m_2)^2} \qquad \beta = \frac{2m_1 m_2}{(m_1 - m_2)}$$

Let $F(x) = \log \dfrac{m(x)}{P(x)}$ where $m(x)$ is the minimum curve as defined in (2.1) for the case $a_1 = -1$, $a_2 = 1$

Then

$$F(x) \equiv \log\left\{\frac{m(x)}{\beta^2(1 - x^2)}\left(\sqrt{\alpha^2 + \beta^2 + 2\alpha\beta x} + \alpha + \beta x\right)\right\}$$

$$F(\pm 1) = 0 \quad \text{and} \quad F''(x) = \frac{d^2}{dx^2}\log\theta(x),$$

where

$$\theta(x) = \sqrt{\alpha^2 + \beta^2 + 2\alpha\beta x} + \alpha + \beta x = \frac{\theta''}{\theta} - \frac{\theta'^2}{\theta^2};$$

since

$$\theta'' = \frac{-(\alpha\beta)^2}{(\alpha^2 + \beta^2 + 2\alpha\beta x)^{3/2}} \quad \text{and} \quad \theta(x) > 0 \text{ in } (-1, 1)$$

$$\frac{d^2}{dx^2}\log\theta(x) < 0 \text{ in } (-1, 1).$$

Hence F is concave downwards and the arc in $(-1, 1)$ lies above the chord, or

$$F(x) \geqq F(\pm 1) = 0$$

or

$$m(x) \geqq P(x).$$

It is clear that in case $m_2 = m_1$ then the parabola of contact will be given by $Y = \frac{m_2}{2}(1 - x^2)$ and coincides with the minimum curve.

PROOF OF THEOREM 3_2: As in theorem 3_1, $a_1 = -1$, $a_2 = 1$ and x_0 the maximum pt. $\neq 0$. Then the parabola

$$Y = \sqrt{\alpha^2 + \beta^2 + 2\alpha\beta x} - \alpha - \beta x$$

where

$$\alpha = \frac{M_0}{2x_0^2} \qquad \beta = -\frac{M_0}{x_0}$$

passes through $(-1, 0)$, $(1, 0)$ and touches $Y = M_0$ at (x_0, M).
Let $M(x)$ be the maximum function defined as in 2.2 and let

$$F(x) = \log \frac{M(x)}{P(x)}.$$

Then $$F(x_0) = F'(x_0) = 0$$

and arguing as in the previous paragraph

$$F''(x) \leqq 0 \qquad \text{in } (a_1 a_2).$$

Hence in $(a_1 a_2)$ $$F(x) \leqq F(x_0) = 0$$

or $$M(x) \leqq P(x)$$

and in case $x_0 = 0$ the parabola will be $Y = M(1 - x^2)$ and coincides with the maximum curve.

UNIVERSITY OF MYSORE
BANGALORE, INDIA

ANNALS OF MATHEMATICS
Vol. 42, No. 4, October, 1941

THE PARTIAL SUMS OF MULTIVALENTLY STAR-LIKE FUNCTIONS

BY M. S. ROBERTSON

(Received September 12, 1940)

1. Introduction

In a recent paper (1) the author has a theorem concerning the n^{th} partial sum of an analytic function univalently star-like[1] with respect to the unit circle. The theorem 7.2 stated that the n^{th} partial sum was also univalently star-like with respect to the circle of radius $1 - 2n^{-1}\log n$. Dr. Otto Szász has kindly pointed out to the author that in the proof of this result an incorrect estimate for $|\rho_n'(z)|$, line 11, page 406, was used. The correct estimate should read

$$|\rho_n'(z)| \leqq r^n(1-r)^{-3}\cdot\{n^2(1-r)^2 + (2n-1)(1-r) + 2\} \tag{1.1}$$

whence line 12 becomes

$$\Re\left[\frac{zf_n'(z)}{f_n(z)}\right] \geqq \frac{1-r}{1+r} - \frac{r^n(1+r)^2}{1-r}\left[\frac{\{(n-1)(1-r)+2\}\{n(1-r)+2\}}{(1-r)^2 - r^n(1+r)^2(1+n-nr)}\right] \tag{1.2}$$

which may be shown to be positive for $n > n_0$ when

$$r = 1 - 4n^{-1}\log n \tag{1.3}$$

if one proceeds by the method used in the proof of Theorem 7.2. Thus Theorem 7.2 of the paper (1) should be corrected to read

THEOREM 7.2: *If*

$$f(z) = z + \sum_2^\infty a_n z^n$$

be regular and univalently star-like with respect to the unit circle then the n^{th} partial sum is univalently star-like for $|z| \leqq 1 - 4n^{-1}\log n$, $n > n_0$.

Whether the constant 4 is the best possible one or not the author is unable to say. There are reasons for belief that the best possible constant is the number 3. For instance, as will be shown in the last section of this paper, the well-known extremal function $z(1-z)^{-2}$ of the theory of univalently star-like functions has a star radius R_n for its n^{th} partial sum with

$$\limsup_{n\to\infty}\left\{\frac{1-R_n}{n^{-1}\log n}\right\} = 3.$$

Since Theorems 7.3, 7.4, and 7.5 of paper (1) were made to depend upon the result of Theorem 7.2, the expression $1 - 2n^{-1}\log n$ in the statements of each of

[1] Multivalently star-like functions of order p are defined in the lines immediately following equation (1.4) of the present paper; univalently star-like functions are those of order 1.

these theorems should be replaced by $1 - 4n^{-1} \log n$. Dr. Szász has informed the author that the sharp estimate $1 - 3n^{-1} \log n$ holds for Theorems 7.4 and 7.5. Instead of making Theorems 7.4 and 7.5 depend upon Theorem 7.2 Dr. Szász obtained his result by showing directly that the n^{th} partial sum of the function $z(1 - 2z \cos \theta + z^2)^{-1}$ is typically-real for $|z| \leqq 1 - 3n^{-1} \log n$. The above function, as was shown in Theorem 7.4, sets the pace for all the functions of the class under consideration. It might be pointed out that the function $z(1 - 2z \cos \theta + z^2)^{-1}$ has its n^{th} partial sum star-like univalently, and à fortiori typically real, for $|z| \leqq 1 - 3n^{-1} \log n$ as may be seen by direct calculation of $\Re[zf_n'(z)/f_n(z)]$ for the n^{th} partial sum $f_n(z)$ of this function. I omit the details but the proof is quite similar to that given for the function $z(1 - z)^{-2}$ in the last section of this paper.

It is now the purpose of this paper, besides correcting the error mentioned above, to generalize Theorem 7.2 to multivalently star-like functions. A definition of multivalently star-like functions follows.

Let

$$f(z) = z^p + a_{p+1}z^{p+1} + \cdots + a_n z^n + \cdots \tag{1.4}$$

be regular and multivalently star-like of order p with respect to the unit circle. This means that within the unit circle $f(z)$ assumes no value more than p times, at least one value p times, and in addition is star-like, i.e.,

$$\Re[zf'(z)/f(z)] > 0, \qquad |z| < 1. \tag{1.5}$$

Geometrically, this means that as z traverses a circle $|z| = r < 1$ in the anti-clockwise direction the radius vector joining the origin to the point $w = f(z)$ in the w-plane turns also in the anti-clockwise direction, completing in this case p revolutions as z traverses the circle once. When $p = 1$ $f(z)$ is said to be univalently star-like. In this case the radius vector cuts the image of $|z| = r < 1$ in the w-plane once only for any direction from the origin.

We have the following generalization of Theorem 7.2 mentioned above.

THEOREM A: *If*

$$f(z) = z^p + a_{p+1}z^{p+1} + \cdots + a_n z^n + \cdots$$

be regular and multivalently star-like of order p with respect to the unit circle then the n^{th} partial sum

$$f_n(z) = z^p + a_{p+1}z^{p+1} + \cdots + a_n z^n$$

for $n > n_0(p)$ is also multivalently star-like of order p for

$$|z| < 1 - (2p + 2)n^{-1} \log n.$$

I do not know if the constant $2p + 2$ can be replaced by a smaller one. However, the constant cannot be replaced by one which is smaller than $2p + 1$. For, as will be shown in the last section of the present paper, the multivalently

star-like function of order p $z^p(1-z)^{-2p}$ has its n^{th} partial sum multivalently star-like of order p in a circle about the origin whose radius is at most R_n where

$$R_n \geqq 1 - (2p+1)n^{-1}\log n$$

$$\limsup_{n\to\infty}\left\{\frac{1-R_n}{n^{-1}\log n}\right\} = 2p+1. \tag{1.6}$$

In the proofs to follow A will denote a constant depending only upon p and the A in one inequality need not necessarily be the same A which occurs again in another or in the same inequality.

2. Preliminary lemmas

Before coming to the main theorem we prove four lemmas.

LEMMA 1. *If*

$$f(z) = z^p + a_{p+1}z^{p+1} + \cdots + a_n z^n + \cdots \tag{2.1}$$

be regular and multivalently star-like of order p for $|z| < 1$, *then*

$$f(z) = [\phi(z)]^p \tag{2.2}$$

where

$$\phi(z) = z + d_2 z^2 + \cdots + d_n z^n + \cdots \tag{2.3}$$

is regular and univalently star-like for $|z| < 1$.

PROOF: Since $f(z)$ is star-like for $|z| < 1$ we have

$$\Re[zf'(z)/f(z)] > 0, \quad |z| < 1. \tag{2.4}$$

Hence

$$p + z\frac{d}{dz}\log\{z^{-p}\cdot f(z)\} = zf'(z)/f(z) = pF(z) \tag{2.5}$$

where $F(z)$ is regular, $F(0) = 1$, $\Re F(z) > 0$ for $|z| < 1$. Thus

$$f(z) = z^p \cdot \exp p\int_0^z z^{-1}\{F(z) - 1\}\,dz \tag{2.6}$$

$$= [\phi(z)]^p$$

where

$$\phi(z) = z\exp\int_0^z z^{-1}\{F(z)-1\}\,dz, \qquad \phi'(0) = 1, \tag{2.7}$$

$$\Re[z\phi'(z)/\phi(z)] = \Re F(z) > 0.$$

Thus $\phi(z)$ is univalently star-like and the proof is complete.

LEMMA 2. *If* $f(z)$ *satisfies the conditions of Lemma 1 then*

$$|f(z)| \geqq r^p(1+r)^{-2p}, \qquad |z| = r < 1, \tag{2.8}$$

$$\Re[zf'(z)/f(z)] \geqq p(1-r)(1+r)^{-1}, \tag{2.9}$$

$$|a_n| \leqq \frac{(n+p-1)!}{(n-p)!\,(2p-1)!}, \qquad n > p. \tag{2.10}$$

PROOF: We have from Lemma 1

$$f(z) = [\phi(z)]^p \tag{2.11}$$

and since $\phi(z)$ is univalent (2) for $|z| < 1$

$$|\phi(z)| \geqq r(1+r)^{-2}, \qquad |z| = r < 1. \tag{2.12}$$

Thus

$$|f(z)| \geqq r^p(1+r)^{-2p}. \tag{2.13}$$

Again, since

$$zf'(z)/f(z) = p \cdot F(z) \tag{2.14}$$

where $\Re F(z) > 0$, $F(0) = 1$, then (3)

$$\Re[zf'(z)/f(z)] = p\Re F(z) \geqq p(1-r)(1+r)^{-1}. \tag{2.15}$$

Finally, it is well known that $\phi(z)$ is majorized (2) by $z(1-z)^{-2}$ since $\phi(z)$ is star-like univalently. Hence $f(z)$ is majorized by $z^p(1-z)^{-2p}$ and thus

$$|a_n| \leqq \frac{(n+p-1)!}{(n-p)!\,(2p-1)!}, \qquad n > p. \tag{2.16}$$

This completes the proof of Lemma 2.

LEMMA 3. *The following identity holds for $r < 1$, p a positive integer.*

$$\sum_{\kappa=n+1}^{\infty} \kappa^p r^\kappa = \frac{r^{n+1}}{(1-r)^{p+1}}\left[p! + \sum_{m=1}^{p} A_m^p (1-r)^m\right], \qquad A_p^p = n^p, \tag{2.17}$$

where A_m^p = a polynomial in n of degree m whose leading coefficient is positive.

PROOF: It is easily verified that

$$\begin{aligned}
\sum_{n+1}^{\infty} \kappa r^\kappa &= \frac{r^{n+1}}{(1-r)^2}[1 + n(1-r)], \\
\sum_{n+1}^{\infty} \kappa^2 r^\kappa &= \frac{r^{n+1}}{(1-r)^3}[2 + (2n-1)(1-r) + n^2(1-r)^2], \\
\sum_{n+1}^{\infty} \kappa^3 r^\kappa &= \frac{r^{n+1}}{(1-r)^4}[6 + (6n-6)(1-r) \\
&\qquad + (3n^2 - 3n + 1)(1-r)^2 + n^3(1-r)^3].
\end{aligned} \tag{2.18}$$

Thus the identity (2.17) is true for $p = 1, 2, 3$. We prove that it is true for any positive integer p by mathematical induction. Assuming (2.17) true for $p = q$ we shall show that in this case it must also be true for $p = q + 1$. By assumption then (2.17) holds with $p = q$. Then if $A_0^q = q!$

$$\sum_{\kappa=n+1}^{\infty} \kappa^{q+1} r^{\kappa} = r \frac{d}{dr} \sum_{m=0}^{q} A_m^q r^{n+1} (1-r)^{m-q-1}$$

$$= r^{n+1}(1-r)^{-q-2} \cdot \sum_{m=0}^{q} A_m^q [(n+m-q)$$

$$\times (1-r)^{m+1} - (m-q-1)(1-r)^m] \tag{2.19}$$

$$= \frac{r^{n+1}}{(1-r)^{q+2}} \Big[(q+1)A_0^q + \sum_{m=0}^{q} \{nA_m^q$$

$$+ (q-m)(A_{m+1}^q - A_m^q)\}(1-r)^{m+1} \Big].$$

In the last summation the coefficient of $(1-r)^{q+1}$ is seen to be n^{q+1} and the coefficient of $(1-r)^{m+1}$ is of degree $(m+1)$ in n since $q-m \geqq 0$ and the leading coefficient of A_m^q is positive for all m. Thus (2.17) holds for $p = q+1$ whenever it is true for $p = q$. But since we know that it holds for $p = 1, 2, 3$ the proof by mathematical induction is complete.

Letting

$$\begin{aligned} f_n(z) &= z^p + a_{p+1}z^{p+1} + \cdots + a_n z^n, \qquad n \geqq p, \\ \rho_n(z) &= a_{n+1}z^{n+1} + a_{n+2}z^{n+2} + \cdots, \\ f(z) &= f_n(z) + \rho_n(z), \end{aligned} \tag{2.20}$$

we proceed to prove

LEMMA 4. *If* $|z| = r = 1 - (2p+2)n^{-1}\log n$ *then for* $n > n_0(p)$ *there is a constant* A *depending upon* p *but not upon* n *for which*

$$\begin{aligned} &|\rho_n(z)| < An^{-2}(\log n)^{-1}, \\ &|\rho'_n(z)| < An^{-1}(\log n)^{-1}, \\ &|f'(z)/f(z)| < An(\log n)^{-1}, \\ &|f(z)| - |\rho_n(z)| \geqq A. \end{aligned} \tag{2.21}$$

PROOF:

$$|\rho_n(z)| \leqq \sum_{\kappa=n+1}^{\infty} \frac{(\kappa+p-1)!}{(\kappa-p)!(2p-1)!} r^{\kappa} < A \sum_{n+1}^{\infty} \kappa^{2p-1} r^{\kappa}$$

$$< Ar^{n+1}(1-r)^{-2p}[(2p-1)! + A_1^{2p-1}(1-r) + \cdots + n^{2p-1}(1-r)^{2p-1}]$$

$$r^{n+1} < \{1-(2p+2)n^{-1}\log n\}^n < e^{-(2p+2)\log n} = n^{-2p-2}, \tag{2.22}$$

$$(2p-1)! + A_1^{2p-1}(1-r) + \cdots + n^{2p-1}(1-r)^{2p-1} < A(\log n)^{2p-1},$$

$$|\rho_n(z)| < An^{-2p-2}\left\{\frac{n}{(2p+2)\log n}\right\}^{2p} \cdot (\log n)^{2p-1} < An^{-2}(\log n)^{-1}.$$

Similarly we have

$$
\begin{aligned}
|\rho_n'(z) &< A\sum_{k=n}^{\infty} k^{2p}r^k \\
&< \frac{Ar^n}{(1-r)^{2p+1}}\,[(2p)! + A_1^{2p}(1-r) + \cdots + (n+1)^{2p}(1-r)^{2p}] \qquad (2.23)\\
&< An^{-2p-2}\left\{\frac{n}{(2p+2)\log n}\right\}^{2p+1}\cdot(\log n)^{2p} < A(n\log n)^{-1}.
\end{aligned}
$$

Again,

$$
\left|\frac{f'(z)}{f(z)}\right| = \frac{p}{r}\,|F(z)| \leqq \frac{p}{r}\left(\frac{1+r}{1-r}\right) < An(\log n)^{-1}, \qquad (2.24)
$$

and

$$
|f(z)| - |\rho_n(z)| \geqq r^p(1+r)^{-2p} - An^{-2}(\log n)^{-1} > A, \qquad (2.25)
$$

for $n > n_0(p)$, $r = 1 - (2p+2)n^{-1}\log n$.

3. Proof of Theorem A

Since

$$
f_n(z) = f(z) - \rho_n(z) \qquad (3.1)
$$

we have for $r = 1 - (2p+2)n^{-1}\log n$

$$
\begin{aligned}
\Re[zf_n'(z)/f_n(z)] &= \Re\left\{[zf'(z)/f(z)] - z\cdot\frac{\rho_n'(z) - \rho_n(z)\cdot f'(z)/f(z)}{f(z) - \rho_n(z)}\right\} \\
&\geqq \Re[zf'(z)/f(z)] - r\cdot\frac{|\rho_n'(z)| + |\rho_n(z)|\cdot|f'(z)/f(z)|}{|f(z)| - |\rho_n(z)|} \qquad (3.2)\\
&\geqq p(1-r)(1+r)^{-1} \\
&\qquad - \frac{A(n\log n)^{-1} + An^{-2}(\log n)^{-1}\cdot n(\log n)^{-1}}{A} \\
&\geqq An^{-1}\log n - A(n\log n)^{-1} > 0, \qquad n > n_0(p).
\end{aligned}
$$

Thus since an harmonic function assumes its minimum value on the boundary, we have for $n > n_0(p)$

$$
\Re[zf_n'(z)/f_n(z)] > 0 \quad \text{for} \quad r \leqq 1 - (2p+2)n^{-1}\log n. \qquad (3.3)
$$

To show that

$$
f_n(z) = z^p + a_{p+1}z^{p+1} + \cdots + a_nz^n
$$

is multivalent of order p in $|z| < 1 - (2p+2)n^{-1}\log n$ we apply a theorem due to S. Ozaki (4) which states that if $f(z)$ is analytic in $|z| \leqq r$ and has p zeros there, none on the circumference, and if for some real α

$$\Re[e^{i\alpha}zf'(z)/f(z)] > 0 \qquad \text{on} \qquad |z| = r,$$

then $f(z)$ is multivalent of order p in the circle $|z| = r$. Applying the theorem to $f_n(z)$ we observe that $f_n(z)$ has no zeros on the boundary of the circle $|z| = 1 - (2p+2)n^{-1}\log n$ since for $n > n_0(p)$

$$|f_n(z)| \geqq |f(z)| - |\rho_n(z)| > A.$$

Further $f_n(z)$ has exactly p zeros inside the given circle. This follows by the theorem of Rouché. On the circle

$$|f_n(z)| > A, \qquad |\rho_n(z)| < An^{-2}(\log n)^{-1}.$$

Therefore on the circle for $n > n_0(p)$ $|f_n(z)| > |\rho_n(z)|$, $f_n(z) \neq 0$. Hence by the theorem of Rouché $f_n(z)$ and $f_n(z) + \rho_n(z) = f(z)$ have the same number of zeros inside the circle. But since $f(z)$ has p zeros at the origin and is multivalent of order p in the unit circle it follows that $f(z)$ has exactly p zeros for $|z| < 1$. We have thus shown that the conditions of the theorem of S. Ozaki hold for $f_n(z)$ (with $\alpha = 0$). Thus $f_n(z)$ is multivalent of order p and star-like by (3.3) for $|z| < 1 - (2p+2)n^{-1}\log n$. This completes the proof of Theorem A.

4. The partial sums of $z^p(1-z)^{-2p}$

Let us consider the following particular multivalently star-like function of order p

$$f(z) = z^p(1-z)^{-2p} = \sum_{k=p}^{\infty} \frac{(k+p-1)!}{(k-p)!(2p-1)!}z^k \tag{4.1}$$

and its n^{th} partial sum

$$f_n(z) = z^p + 2pz^{p+1} + \cdots + \frac{(n+p-1)!}{(n-p)!(2p-1)!}z^n. \tag{4.2}$$

We define R_n to be the radius of the largest circle $|z| = R_n$ within which $f_n(z)$ is multivalently star-like of order p. We shall show that for $n > n_0(p)$

$$R_n \geqq 1 - (2p+1)n^{-1}\log n$$

$$\limsup_{n\to\infty}\left\{\frac{1-R_n}{n^{-1}\log n}\right\} = 2p+1. \tag{4.3}$$

We shall first establish by mathematical induction the following identity

$$f_n(z) = z^p(1-z)^{-2p}\left[1 - \frac{(n+p)!}{(n-p)!(2p-1)!}\int_0^z z^{n-p}(1-z)^{2p-1}\,dz\right] \tag{4.4}$$

The above formula is easily verified to be true for $n = p$. Let us now assume it correct for a value n and then show that this implies that the formula is also correct for the value $n + 1$. We first note an integration by parts gives

$$\frac{2p}{n-p+1}\int_0^z z^{n-p+1}(1-z)^{2p-1}\,dz$$

$$= -\frac{(1-z)^{2p}z^{n-p+1}}{n-p+1} + \int_0^z (1-z)(1-z)^{2p-1}z^{n-p}\,dz$$

$$= -\frac{(1-z)^{2p}z^{n-p+1}}{n-p+1} + \int_0^z z^{n-p}(1-z)^{2p-1}\,dz - \int_0^z z^{n-p+1}(1-z)^{2p-1}\,dz,$$

$$\frac{n+p+1}{n-p+1}\int_0^z z^{n-p+1}(1-z)^{2p-1}\,dz = \int_0^z z^{n-p}(1-z)^{2p-1}\,dz - \frac{z^{n-p+1}(1-z)^{2p}}{n-p+1}.$$

Using the last equality we obtain

$$f_{n+1}(z) = f_n(z) + \frac{(n+p)!}{(n-p+1)!(2p-1)!}z^{n+1}$$

$$= z^p(1-z)^{-2p}\left[1 - \frac{(n+p+1)!}{(n-p+1)!(2p-1)!}\right.$$

$$\left.\cdot\frac{n-p+1}{n+p+1}\left\{\int_0^z z^{n-p}(1-z)^{2p-1}\,dz - \frac{z^{n-p+1}(1-z)^{2p}}{n-p+1}\right\}\right]$$

$$= z^p(1-z)^{-2p}\left[1 - \frac{(n+p+1)!}{(n-p+1)!(2p-1)!}\int_0^z z^{n-p+1}(1-z)^{2p-1}\,dz\right].$$

Thus we have shown that the formula is correct for the value $n + 1$. This completes the proof of the formula.

Next we differentiate logarithmically the formula (4.4) and obtain

$$zf_n'(z)/f_n(z) = p\left(\frac{1+z}{1-z}\right)$$

$$- \frac{(n+p)!z^{n-p+1}(1-z)^{2p-1}}{\left\{(2p-1)!(n-p)! - (n+p)!\int_0^z z^{n-p}(1-z)^{2p-1}\,dz\right\}}. \tag{4.5}$$

For value of $|z| = r < 1$ for which

$$(2p-1)!(n-p)! - (n+p)!\int_0^r r^{n-p}(1+r)^{2p-1}\,dr > 0$$

we may write

$$\Re[zf_n'(z)/f_n(z)] \geqq p\left(\frac{1-r}{1+r}\right)$$

$$- \frac{(n+p)!r^{n-p+1}(1+r)^{2p-1}}{(2p-1)!(n-p)! - (n+p)!\int_0^r r^{n-p}(1+r)^{2p-1}\,dr}. \tag{4.6}$$

Since $\int_0^z z^{n-p}(1-z)^{2p-1}\,dz$ is a polynomial with real coefficients whose signs alternate, it is seen that for $(n-p)$ odd equality is obtained in (4.6) when $z=-r$.

We shall then have

$$\Re[zf_n'(z)/f(z)] > 0$$

provided

(i) $$(2p-1)!(n-p)!p(1-r) - (n+p)!p(1-r)\int_0^r r^{n-p}(1+r)^{2p-1}\,dr - (n+p)!r^{n-p+1}(1+r)^{2p} > 0$$

and

(ii) $$(2p-1)!(n-p)! - (n+p)!\int_0^r r^{n+p}(1+r)^{2p-1}\,dr > 0.$$

It is readily seen that the inequality (i) implies the inequality (ii).

Let the left-hand side of the inequality (i) be denoted by $T_n(r)$ and let $r = 1-\frac{\alpha}{n}$ where $\alpha = (2p+1)\log n$.

For $n > n_0(p)$

$$T_n(r) > (n-p)!(2p)!\tfrac{1}{2}(1-r) - (n+p)!p(1-r)r^{n-p+1}2^{2p-1} - (n+p)!2^{2p}r^{n-p+1}$$

$$\frac{T_n(r)}{(n-p)!} > (2p+1)!(2n)^{-1}\log n - (2n)^{2p}\cdot p\cdot 1\cdot\left(1-\frac{\alpha}{n}\right)^{1-p}e^{-\alpha}\cdot 2^{2p-1} - (2n)^{2p}\cdot 2^{2p}\left(1-\frac{\alpha}{n}\right)^{1-p}e^{-\alpha}$$

$$> (2p+1)!(2n)^{-1}\log n - A(p)n^{2p}e^{-\alpha}$$

$$\geqq (2p+1)!(2n)^{-1}\log n - A(p)n^{-1}$$

where $A(p)$ is a constant depending upon p. Hence for $n > n_0(p)$

(4.7) $$T_n\{1-(2p+1)n^{-1}\log n\} > 0.$$

On the other hand if we denote by $A(p,\epsilon)$ a constant depending only upon p and ϵ, $0<\epsilon<1$, then for $n > n_0(p,\epsilon)$ when $r = 1-(2p+1-\epsilon)n^{-1}\log n$ we have

$$\frac{T_n(r)}{(n-p)!} < (2p)!(2p+1-\epsilon)(2n)^{-1}\log n - \frac{(n+p)!r^n}{(n-p)!}$$

$$< (2p)!(2p+1-\epsilon)(2n)^{-1}\log n - n^{2p}\cdot A(p,\epsilon)n^{\epsilon-1-2p}$$

$$\leqq (2p)!(2p+1-\epsilon)(2n)^{-1}\log n - A(p,\epsilon)n^{\epsilon-1}.$$

Hence for $n > n_0(p, \epsilon)$

$$T_n\{1 - (2p + 1 - \epsilon)n^{-1} \log n\} < 0. \tag{4.8}$$

Thus we have shown that for $|z| \leqq 1 - (2p + 1)n^{-1} \log n$ and for $n > n_0(p)$

$$\Re[zf_n'(z)/f_n(z)] > 0 \tag{4.9}$$

and that the constant $(2p + 1)$ cannot be replaced by a smaller one for all $(n - p)$ odd.

A proof similar to that given at the end of the preceding section shows that the n^{th} partial sum for the power series of $z^p(1 - z)^{-2p}$ is multivalent of order p for $|z| < 1 - (2p + 1)n^{-1} \log n$. Thus we have for $n > n_0(p)$

$$R_n \geqq 1 - (2p + 1)n^{-1} \log n, \tag{4.10}$$

$$\limsup_{n \to \infty} \left\{ \frac{1 - R_n}{n^{-1} \log n} \right\} = 2p + 1. \tag{4.11}$$

RUTGERS UNIVERSITY,
NEW BRUNSWICK, N. J.

REFERENCES

1. M. S. Robertson, *On the theory of univalent functions*, Annals of Mathematics, vol. 37, 1936, pp. 374–408.
2. L. Bieberbach, *Lehrbuch der Funktionentheorie*, Berlin, vol. 2, 1931, pp. 71–83.
3. G. Pólya and G. Szegö, *Aufgaben und Lehrsatze aus der Analysis*, Berlin, vol. 1, 1925, p. 140.
4. S. Ozaki, *Some remarks on the univalency and multivalency of functions*, Science Reports of Tokyo Bunrika Daigaku, sect. A, 2, no. 32, 1934, pp. 41–55.

ANNALS OF MATHEMATICS
Vol. 42, No. 4, October, 1941

TWO-SIDED IDEALS AND CONGRUENCES IN THE RING OF BOUNDED OPERATORS IN HILBERT SPACE

J. W. CALKIN

(Received August 29, 1940)

Introduction

The developments of the present paper center around the observation that the ring $\mathfrak{B}$ of bounded everywhere defined operators in Hilbert space contains non-trivial two-sided ideals.[1] This fact, which has escaped all but oblique notice in the development of the theory of operators, is of course fundamental from the point of view of algebra and at the same time differentiates $\mathfrak{B}$ sharply from the ring of all linear operators over a unitary space with finite dimension number.

As examples of two-sided ideals in $\mathfrak{B}$ we may mention here the class of all operators A such that $\mathfrak{R}(A)$, the range of A, has a finite dimension number, the class of all operators of Hilbert-Schmidt type,[2] and the class $\mathfrak{T}$ of all totally continuous operators. Except for the ideal (0), every two-sided ideal in $\mathfrak{B}$ contains the first ideal mentioned, and except for the ideal $\mathfrak{B}$ itself, every two-sided ideal in $\mathfrak{B}$ is contained in the ideal $\mathfrak{T}$. Moreover, on the basis of the special spectral properties of the self-adjoint members of $\mathfrak{T}$, it is possible to characterize every two-sided ideal in $\mathfrak{B}$ very simply in terms of the spectra of its nonnegative self-adjoint elements; for both the formulation and the proof of this result, which together with the facts mentioned above is discussed in §1, the author is indebted to J. v. Neumann.

The restriction of our attention to those ideals in $\mathfrak{B}$ which are two-sided is basic for the points which we wish to develop; the two-sidedness compensates for the absence of commutativity in $\mathfrak{B}$ in such a way as to permit the construction of quotient rings by the standard methods of abstract algebra.[3] These rings, which are of course homomorphs of $\mathfrak{B}$ with respect to addition and multiplication, are also homomorphs of $\mathfrak{B}$ with respect to the operation *, and exhibit all of the formal properties of matrix algebras. This is established in §2, and there also various properties of the associated congruences in $\mathfrak{B}$ are discussed.

The remainder of the paper deals solely with the quotient ring $\mathfrak{B}/\mathfrak{T}$, where $\mathfrak{T}$ is the ideal of totally continuous operators, and the associated congruence in $\mathfrak{B}$. For essentially topological reasons, this is the only one of the quotient rings in

[1] An additive subset $\mathfrak{J}$ of $\mathfrak{B}$ is a left (right) ideal if it contains AB (BA) for all A in $\mathfrak{J}$ and B in $\mathfrak{B}$. If $\mathfrak{J}$ is both a left and right ideal, it is called two-sided. See references [1] and [19] at the end of the paper for the elementary properties of ideals.

[2] For a discussion in abstract terms of operators of this type (operators of "finite norm"), see reference [17], pp. 65–70.

[3] See, for example, [1], pp. 252–253, [19], pp. 56–57.

question which at present appears susceptible of deep analysis. For while there is in general no apparent way of introducing a topology in quotient rings over $\mathfrak{B}$, the ring $\mathfrak{B}/\mathfrak{T}$ is actually a complete metric space, the norm deriving very simply from the norm (i.e., bound) in $\mathfrak{B}$ itself (§3).

Moreover, it is even possible to interpret $\mathfrak{B}/\mathfrak{T}$ as an algebraic ring[4] of operators in a suitably defined complex Euclidean space $\mathfrak{L}$ whose dimension number is the cardinal number of the continuum.[5] Or, to put it differently, there exists in the ring of bounded everywhere-defined operators over $\mathfrak{L}$ a subset $\mathfrak{M}$ which is a $(+, \cdot, *)$-isomorphism of $\mathfrak{B}/\mathfrak{T}$. Furthermore, this correspondence is even an isometry, the norm of an element of $\mathfrak{B}/\mathfrak{T}$ being the bound of the corresponding element of $\mathfrak{M}$. These facts are all established in §§4, 5, and in §5, various properties of the algebraic ring $\mathfrak{M}$ are discussed.

Apart from its intrinsic interest, the analysis of the ring $\mathfrak{M}$ yields in a very simple way theorems of considerable depth concerning $\mathfrak{B}$ (§5). Of results of this sort, we shall here mention only a generalization of Weyl's classical theorem comparing the spectra of two self-adjoint operators with totally continuous difference [20].[6]

Before proceeding, we should like to point out various further developments which are suggested by the present paper and to which we propose to return at another time.

First, while the algebraic ring $\mathfrak{M}$ is not closed with respect to the weak topology for operators and so is not an operator ring in the sense of von Neumann [9, 11] one obtains such a ring by adjoining to $\mathfrak{M}$ its weak condensation points in the ring of bounded everywhere-defined operators in $\mathfrak{L}$. This "closure" $\mathbf{R}(\mathfrak{M})$ is of considerable interest from the point of view of the Murray-von Neumann theory of factors, especially since various preliminary results concerning $\mathbf{R}(\mathfrak{M})$ and $\mathfrak{M}'$ suggest that they may be factors of class III_∞.[7]

Second, while the factors of class II_1 and III_∞ of Murray and von Neumann are like those of class I_n, $n < \aleph_0$, simple rings, those of class II_∞ are not. The construction of quotient rings over such a factor is therefore possible and gives rise to various questions analogous to those concerning $\mathfrak{B}$ with which we deal here.[8]

Finally, we should like to point out that the maximal property of the ideal $\mathfrak{T}$ of totally continuous operators described above does not persist if one considers

---

[4] We use the term algebraic ring with reference to operators to denote a class closed with respect to the operations $+, \cdot, *$, and scalar multiplication. This is not a ring in the sense of von Neumann, [11], since no topological conditions are imposed.

[5] For the theory of complex Euclidean spaces of arbitrary dimension number, see references [6], [7], [13], [16].

[6] All numbers in brackets refer to the bibliography at the end of the paper.

[7] For the notion of factors, and their classification, see [9]; for a construction of factors of class III_∞, see [15].

[8] The class of members of a II_∞ factor which are "normed" in the sense of [15] is a two-sided ideal.

instead of Hilbert space, a space $\mathfrak{H}$ with dimension number greater than $\aleph_0$. For example, in such a space, the class of all operators A such that $\mathfrak{R}(A)$ contains no closed subspace with dimension number exceeding $\aleph_0$ is a non-trivial ideal. Moreover, if m is the dimension number of $\mathfrak{H}$, $m \geqq \aleph_0$, the class of all operators A which have the property that $\mathfrak{R}(A)$ contains no closed linear subspace of $\mathfrak{H}$ with dimension number m is a two-sided ideal different from $\mathfrak{B}$ and is identical with $\mathfrak{T}$ if $m = \aleph_0$. Furthermore, in every case, this ideal is maximal in the same sense in which $\mathfrak{T}$ is for $m = \aleph_0$, and as von Neumann has pointed out to the author, the spectral characterization of ideals can be extended in a satisfactory way to describe the general case. Finally, investigations which are still incomplete suggest that a considerable portion of the analysis of the present paper has a counterpart in every case.

1. Two-sided ideals in $\mathfrak{B}$

We proceed now to the characterization of all two-sided ideals in $\mathfrak{B}$.

THEOREM 1.1. *If $\mathfrak{I}$ is a left (right) ideal, the set $\mathfrak{I}^*$ of all adjoints A^* of elements A of $\mathfrak{I}$ is a right (left) ideal.*

Since $\mathfrak{I}^*$ is evidently closed with respect to addition, it is necessary only to show—when $\mathfrak{I}$ is a left ideal—that $\mathfrak{I}^*$ contains BA^* for all B in $\mathfrak{B}$ and A in $\mathfrak{I}$. But this follows at once from the relation $BA^* = (AB^*)^*$ and the fact that AB^* belongs to $\mathfrak{I}$ along with A. When $\mathfrak{I}$ is a right ideal, an analogous argument is valid.

THEOREM 1.2. *A necessary and sufficient condition that a left (right) ideal $\mathfrak{I}$ be two-sided is that $\mathfrak{I} = \mathfrak{I}^*$.*

The sufficiency of the condition is an immediate consequence of Theorem 1.1. Now suppose $\mathfrak{I}$ is two-sided, A an arbitrary member of $\mathfrak{I}$, $A = WB$ its canonical decomposition.[9] Then $A^* = BW^* = W^*AW^*$ is clearly in $\mathfrak{I}$ and $\mathfrak{I} = \mathfrak{I}^*$.

Since we now have no further direct concern with left or right ideals we shall refer to two-sided ideals merely as ideals.

THEOREM 1.3. *The class $\mathfrak{T}$ of all totally continuous operators in $\mathfrak{B}$ is an ideal.*

That $\mathfrak{T}$ is an additive class is obvious from the definition of a totally continuous operator; T is totally continuous if it takes every bounded set into a compact set. Moreover, by a well-known theorem, $\mathfrak{T} = \mathfrak{T}^*$.[10] Finally, since every member A of $\mathfrak{B}$ clearly takes bounded sets into bounded sets, $\mathfrak{T}$ is a left ideal. Hence by Theorem 2 it is an ideal.

Throughout the remainder of the paper $\mathfrak{T}$ has the same meaning as in Theorem 1.3.

THEOREM 1.4. *Let $\mathfrak{I}$ be an arbitrary ideal in $\mathfrak{B}$. Then either $\mathfrak{I} = \mathfrak{B}$ or $\mathfrak{I} \subseteq \mathfrak{T}$.*

The proof of this theorem is based on a characteristic property of totally continuous operators which the writer has noted elsewhere.[11] According to

---

[9] For the notion of canonical decomposition, see [9] and [12].

[10] [2], p. 100, Théorème 4.

[11] [3], Lemma 3.1.

this result a member T of $\mathfrak{B}$ is totally continuous if and only if every closed linear manifold in its range has a finite dimension number. Hence, if $\mathfrak{I}$ is an ideal which is not contained in $\mathcal{T}$, $\mathfrak{I}$ contains an element A such that $\mathfrak{R}(A)$ contains a Hilbert space $\mathfrak{M}$. We denote by $\mathfrak{N}$ the manifold of zeros of A and by A_1 the transformation induced on $\mathfrak{H} \ominus \mathfrak{N}$ by A.[12] The transformation A_1 evidently possesses an inverse and the same range as A; moreover the fact that A_1 is bounded assures us that the set $\mathfrak{A} = A_1^{-1}\mathfrak{M}$ is closed. Thus $\mathfrak{A}$ is also a Hilbert space. Hence there exist in $\mathfrak{B}$ partially isometric operators[13] X, Y both with initial sets $\mathfrak{H}$, while $\mathfrak{R}(X) = \mathfrak{A}$, $\mathfrak{R}(Y) = \mathfrak{M}$. Therefore the operator $B = Y^*AX$ has domain and range identically $\mathfrak{H}$ and belongs to $\mathfrak{B}$. But $Bf = 0$ implies either $Xf = 0$, Xf in $\mathfrak{N}$, or AXf in $\mathfrak{H} \ominus \mathfrak{M}$, and all of these are impossible in view of the definition of X and Y. Hence B^{-1} exists and since it is closed with domain $\mathfrak{H}$, it belongs to $\mathfrak{B}$. Therefore $I = B^{-1}B$ belongs to $\mathfrak{I}$ and $\mathfrak{I} = \mathfrak{B}$.

THEOREM 1.5. *Let $\mathfrak{I}$ be an arbitrary ideal in $\mathfrak{B}$, $\mathfrak{I}_0$ the class of nonnegative definite self-adjoint transformations in $\mathfrak{I}$. Then $\mathfrak{I}_0$ is the class of all operators $B = (A^*A)^{\frac{1}{2}}$ such that A is an element of $\mathfrak{I}$, and if $\mathfrak{I}_1$ is an ideal containing $\mathfrak{I}_0$, then $\mathfrak{I}_1 \supseteq \mathfrak{I}$.*

If B is in $\mathfrak{I}_0$, it is obvious that B is of the form described in the theorem. On the other hand, if A is an arbitrary element of $\mathfrak{I}$, $(A^*A)^{\frac{1}{2}} = B = W^*A$, where W is partially isometric; hence $(A^*A)^{\frac{1}{2}}$ belongs to $\mathfrak{I}$, and thus to $\mathfrak{I}_0$. Finally the relations $A = WB$, $B = (A^*A)^{\frac{1}{2}}$ assure us that $\mathfrak{I}$ is the smallest ideal containing $\mathfrak{I}_0$.

If $\mathfrak{I}$ is an ideal in $\mathfrak{B}$, we call the subset $\mathfrak{I}_0$ defined in Theorem 1.5 the positive part of $\mathfrak{I}$. In order to characterize those subsets of the positive part of $\mathfrak{B}$ which appear as the positive parts of ideals, we recall that every self-adjoint operator T in $\mathcal{T}$ can be reduced to diagonal form; that is, for each such T there exists a complete orthonormal set $\{\varphi_n\}$ in $\mathfrak{H}$ and a sequence $\{\lambda_n\}$ such that $T\varphi_n = \lambda_n\varphi_n$, $n = 1, 2, \cdots$. Moreover, the sequence $\{\lambda_n\}$ is convergent to zero, and nonnegative if T is in the positive part of $\mathcal{T}$. We call this sequence a characteristic sequence of T.

We next observe that if T belongs to the positive part of an ideal $\mathfrak{I}$ and $\{\psi_n\}$ is another complete orthonormal set in $\mathfrak{H}$, then the member A of $\mathcal{T}$ defined by the equations $A\psi_n = \lambda_n\psi_n$, $n = 1, 2, \cdots$, also belongs to the positive part of $\mathfrak{I}$, since $A = UTU^{-1}$, where U is unitary. Hence, if $\mathfrak{I}$ is an arbitrary ideal in $\mathfrak{B}$, $\mathfrak{I} \subseteq \mathcal{T}$ the set of all characteristic sequences $\{\lambda_n\}$ belonging to members of $\mathfrak{I}_0$ may without ambiguity be called the *spectral set* of $\mathfrak{I}$.

We now characterize intrinsically those subsets of the class of nonnegative sequences with limit zero which occur as spectral sets. As is indicated in the introduction, this result is due to J. v. Neumann.

DEFINITION 1.1. *Let $\mathfrak{T}$ denote the class of all infinite sequences of nonnegative*

[12] We regard A as a transformation between Hilbert spaces in the sense of [8].

[13] For this concept, see [9], Definition 4.3.1.

numbers which converge to zero. A subset $\mathfrak{J}$ of $\mathfrak{T}$ is called an ideal set if it has the following properties:

(i) *If $\{\lambda_n\}$ is in $\mathfrak{J}$ and π denotes an arbitrary permutation of the positive integers, $\{\lambda_{\pi(n)}\}$ is in $\mathfrak{J}$;*

(ii) *if $\{\lambda_n\}$ and $\{\mu_n\}$ are in $\mathfrak{J}$, so is $\{\lambda_n + \mu_n\}$;*

(iii) *if $\{\lambda_n\}$ is in $\mathfrak{J}$, $\{\mu_n\}$ in $\mathfrak{T}$ and the inequality $\lambda_n \geqq \mu_n$ holds for all n, then $\{\mu_n\}$ is in $\mathfrak{J}$.*

Our object now is to prove that every spectral set is an ideal set, and conversely. We require first two lemmas concerning ideal sets.

LEMMA 1.1. *Let $\mathfrak{J}$ be an ideal set, $\{\lambda_n\}$ an element of $\mathfrak{J}$ with infinitely many terms different from zero. Then the subsequence of positive terms of $\{\lambda_n\}$ belongs to $\mathfrak{J}$.*

We distinguish two cases according as $\{\lambda_n\}$ contains a finite or an infinite number of zeros. In view of condition (i) of Definition 1, we can assume in the first case $\lambda_1 = \lambda_2 = \cdots = \lambda_N = 0$ and $0 < \lambda_{n+1} \leqq \lambda_n$, $n = N+1, N+2, \cdots$. We then have to show that $\{\lambda_{N+n}\}$ belongs to $\mathfrak{J}$. Let $\{\mu_n\}$ be the sequence defined by the equations

$$\mu_n = \lambda_{N+n}, \qquad n = 1, 2, \cdots, N,$$
$$\mu_n = 0, \qquad n = N+1, N+2, \cdots.$$

Then $\{\mu_n\}$ is dominated by a permutation of $\{\lambda_n\}$ and hence belongs to $\mathfrak{J}$, by conditions (i) and (iii) of Definition 1. Moreover, by condition (ii), $\{\lambda_n + \mu_n\}$ belongs to $\mathfrak{J}$. But we have $\lambda_{N+n} \leqq \lambda_n + \mu_n$, $n = 1, 2, \cdots$, and hence $\{\lambda_{N+n}\}$ belongs to $\mathfrak{J}$.

Now suppose $\{\lambda_n\}$ contains an infinite number of zeros. Again invoking (i), we assume

$$\lambda_{2k} \geqq \lambda_{2(k+1)} > 0, \qquad \lambda_{2k-1} = 0, \qquad k = 1, 2, \cdots.$$

We must then show that $\{\lambda_{2n}\}$ belongs to $\mathfrak{J}$. We define a sequence $\{\mu_n\}$ by the equations

$$\mu_{2k} = \lambda_{2k-1}, \qquad \mu_{2k-1} = \lambda_{2k}, \qquad k = 1, 2, \cdots.$$

Then $\{\mu_n\}$ is a permutation of $\{\lambda_n\}$ and hence belongs to $\mathfrak{J}$. But then $\{\lambda_n + \mu_n\}$ belongs to $\mathfrak{J}$, and we have

$$\lambda_{2n} \leqq \lambda_n + \mu_n$$

Thus, by (iii), $\{\lambda_{2n}\}$ belongs to $\mathfrak{J}$ as we wished to prove.

LEMMA 1.2. *Let $\mathfrak{J}$ be an ideal set, $\{\lambda_n\}$ an element of $\mathfrak{J}$ consisting solely of positive terms. Then any sequence $\{\mu_n\}$ which contains $\{\lambda_n\}$ as a subsequence and which, except for this subsequence, consists solely of zeros, is a member of $\mathfrak{J}$.*

We define

$$\lambda^{(1)}_{2k-1} = 0, \qquad \lambda^{(1)}_{2k} = \lambda_{2k}, \qquad k = 1, 2, \cdots,$$
$$\lambda^{(2)}_{2k-1} = \lambda_{2k-1}, \qquad \lambda^{(2)}_{2k} = 0, \qquad k = 1, 2, \cdots.$$

Then $\{\lambda_n^{(1)}\}$ and $\{\lambda_n^{(2)}\}$ are both dominated by $\{\lambda_n\}$ and therefore belong to $\mathfrak{J}$. Letting j_k denote the integer immediately preceding $k/3$, we now define two permutations π_1, π_2 by the equations,

$$\pi_1(k + j_k) = 2k - 1, \qquad \pi_1(4k) = 2k, \qquad k = 1, 2, \cdots,$$

$$\pi_2(4k - 2) = 2k - 1, \qquad \pi_2(k + j_{k+2}) = 2k, \qquad k = 1, 2, \cdots.$$

The effect of the first of these permutations is to take the set of all integers divisible by 4 into the set of all even integers, preserving order; and to take the set of all integers not divisible by 4 into the set of all odd integers, again preserving order. The effect of the second is to take all even integers not divisible by 4 into the set of all odd integers, and the complementary set of integers into the set of all even integers, order being preserved in both cases. Hence the sequence

$$\{\nu_n\} = \{\lambda^{(1)}_{\pi_1(n)} + \lambda^{(2)}_{\pi_2(n)}\}$$

which clearly belongs to $\mathfrak{J}$, is related to the sequence $\{\lambda_n\}$ by the equations

$$\nu_{2k} = \lambda_k, \qquad \nu_{2k-1} = 0, \qquad k = 1, 2, \cdots.$$

But any sequence $\{\mu_n\}$ derived from $\{\lambda_n\}$ in the manner described in the theorem, and containing an infinite number of zero terms, is a permutation of $\{\nu_n\}$ and therefore belongs to $\mathfrak{J}$.

Thus, to complete the proof, we have only to dispose of the case that $\{\mu_n\}$ contains only a finite number of zero terms. In this case it is convenient to assume that $\{\lambda_n\}$ is monotone. We can then complete the proof by showing that the sequence $\{\mu_n\}$ defined by the equations

$$\mu_n = 0, \qquad n = 1, 2, \cdots, N - 1, \qquad \mu_{N+n-1} = \lambda_n, \qquad n = 1, 2, \cdots,$$

belongs to $\mathfrak{J}$. To do this we set

$$\nu_n = \lambda_1, \qquad n = 1, 2, \cdots, N, \qquad \nu_n = \lambda_n \qquad n = N + 1,$$

$$N + 2, \cdots, 2N, \cdots.$$

Then, from the validity of the lemma in the case of an infinite number of zeros and property (ii) of ideal sets we can conclude that $\{\nu_n\}$ is in $\mathfrak{J}$. But, since we clearly have

$$\mu_n \leqq \nu_n$$

it follows that $\{\mu_n\}$ is in $\mathfrak{J}$.

THEOREM 1.6. *Let $\mathcal{I}$ be an ideal in $\mathfrak{B}$, $\mathcal{I} \subseteq \mathcal{T}$, and let $\mathfrak{J}$ be its spectral set. Then $\mathfrak{J}$ is an ideal set. Conversely, if $\mathfrak{J}$ is an arbitrary ideal set in $\mathfrak{T}$, there exists an ideal in $\mathfrak{B}$ whose spectral set is $\mathfrak{J}$. This correspondence between ideals in $\mathfrak{B}$ and ideal sets in $\mathfrak{T}$ is an isomorphism with respect to the relation $\subset$.*

If $\mathcal{I}$ is a two-sided ideal in $\mathfrak{B}$, $\mathcal{I} \subseteq \mathcal{T}$, we consider an arbitrary member A of $\mathcal{I}_0$ in diagonal form; $A\varphi_n = \lambda_n\varphi_n$, $n = 1, 2, \cdots$, where $\{\varphi_n\}$ is a complete orthonormal set in $\mathfrak{H}$. Then $\{\lambda_n\}$ is in the spectral set $\mathfrak{J}$ of $\mathcal{I}$, and since A can also

be put in diagonal form with reference to any permutation of the sequence $\{\varphi_n\}$, all permutations of $\{\lambda_n\}$ clearly belong to $\mathfrak{J}$ too.

Now let B be any other member of $\mathfrak{I}_0$, also in diagonal form; $B\psi_n = \mu_n\psi_n$, $n = 1, 2, \cdots$, and let U be the unitary operator defined on $\{\varphi_n\}$ by the equations $U\varphi_n = \psi_n$, $n = 1, 2, \cdots$. Then $A + U^{-1}BU$ is in $\mathfrak{I}_0$ and

$$(A + U^{-1}BU)\varphi_n = (\lambda_n + \mu_n)\varphi_n, \qquad n = 1, 2, \cdots.$$

Thus $\{\lambda_n + \mu_n\}$ belongs to $\mathfrak{J}$.

Finally, let $\{\mu_n\}$ be a sequence such that $\mu_n \leqq \lambda_n$, $n = 1, 2, \cdots$. We define an operator B_0 on $\{\varphi_n\}$ by the equations

$$B_0\varphi_n = (\mu_n/\lambda_n)\varphi_n, \qquad \lambda_n \neq 0,$$
$$B_0\varphi_n = 0 \qquad , \qquad \lambda_n = 0.$$

Then B_0 has a closed linear extension B which belongs to $\mathfrak{B}$. Thus AB belongs to $\mathfrak{I}$. But $AB\varphi_n = \mu_n\varphi_n$, $n = 1, 2, \cdots$, and therefore AB belongs to $\mathfrak{I}_0$, $\{\mu_n\}$ to $\mathfrak{J}$.

We turn now to the converse part of the theorem, denoting by $\mathfrak{J}$ an arbitrary ideal set in $\mathfrak{T}$. We designate by $\mathfrak{I}$ the set of all totally continuous operators A such that a characteristic sequence of $(A^*A)^{\frac{1}{2}}$ belongs to $\mathfrak{J}$. Since, if $\mathfrak{I}$ is an ideal, $\mathfrak{J}$ is obviously its spectral set, we have only to show that $\mathfrak{I}$ is an ideal.

We begin by considering an arbitrary pair of operators A and B of $\mathfrak{I}$, with the object of showing that $A + B$ belongs to $\mathfrak{I}$. To this end we consider the operators

$$D_1 = (A^*A)^{\frac{1}{2}}, \qquad D_2 = (B^*B)^{\frac{1}{2}}, \qquad D = (A^* + B^*)(A + B)^{\frac{1}{2}}$$

and characteristic sequences $\{\lambda_n\}$, $\{\mu_n\}$, $\{\nu_n\}$ of D_1, D_2, D, respectively. We have then to show that $\{\nu_n\}$ belongs to $\mathfrak{J}$.

For convenience, we assume that the positive terms of each sequence are arranged in montone order; in addition, if any one of the sequences contains only a finite number of positive terms, we assume that the sequence itself is monotone. This assumption is clearly not restrictive, in view of condition (i) of Definition 1. Now let $\{\lambda_n'\}$ be identical with the subsequence of positive terms of $\{\lambda_n\}$ if that subsequence is infinite, identical with $\{\lambda_n\}$ itself in the alternative case, and let $\{\mu_n'\}$ and $\{\nu_n'\}$ be defined in the same way with reference to $\{\mu_n\}$ and $\{\nu_n\}$ respectively. Then, by Lemma 1.1, $\{\lambda_n'\}$ and $\{\mu_n'\}$ belong to $\mathfrak{J}$; and by Lemma 1.2, $\{\nu_n\}$ belongs to $\mathfrak{J}$ if $\{\nu_n'\}$ does. Thus we can prove that $A + B$ belongs to $\mathfrak{I}$ by showing that $\{\nu_n'\}$ belongs to $\mathfrak{J}$.

To establish the latter result, we note first the relation

$$2(A^*A + B^*B) - (A^* + B^*)(A + B) = (A^* - B^*)(A - B),$$

which implies that $2(D_1^2 + D_2^2) - D^2$ is nonnegative definite. Holding this fact in reserve, we then recall the theorem of Courant [4],[14] characterizing the

[14] The paper cited deals only with integral operators with continuous kernels. However, the more general result required here is readily obtained.

sequence $\{\nu_n'^2\}$ associated with the operator D^2 in terms of maxima and minima. For our purposes, we may state this result as follows:

$$\nu_n'^2 = \operatorname*{Min}_{\dim(\mathfrak{M})\leqq n-1} \operatorname*{Max}_{\substack{f\in\mathfrak{H}\ominus\mathfrak{M}\\ |f|=1}} (D^2 f, f) \tag{1}$$

where $\dim(\mathfrak{M})$ is the dimension number of $\mathfrak{M}$. Similarly,

$$\lambda_n'^2 = \operatorname*{Min}_{\dim(\mathfrak{N})\leqq n-1} \operatorname*{Max}_{\substack{f\in\mathfrak{H}\ominus\mathfrak{N}\\ |f|=1}} (D_1^2 f, f), \tag{2}$$

$$\mu_n'^2 = \operatorname*{Min}_{\dim(\mathfrak{P})\leqq n-1} \operatorname*{Max}_{\substack{f\in\mathfrak{H}\ominus\mathfrak{P}\\ |f|=1}} (D_2^2 f, f). \tag{3}$$

On the basis of these relations we are able to conclude that the inequality

$$\nu_n'^2 \leqq 2(\lambda_j'^2 + \mu_k'^2) \tag{4}$$

is valid provided we have $j + k \leqq n + 1$. For if $\nu_n'^2 > 2(\lambda_j'^2 + \mu_k'^2)$ holds, we have, in view of (2) and (3),

$$\nu_n'^2/2 > \operatorname*{Min}_{\dim(\mathfrak{N})\leqq j-1} \operatorname*{Max}_{f\in\mathfrak{H}\ominus\mathfrak{N}} (D_1^2 f, f) + \operatorname*{Min}_{\dim(\mathfrak{P})\leqq k-1} \operatorname*{Max}_{f\in\mathfrak{H}\ominus\mathfrak{P}} (D_2^2 f, f)$$

and this implies

$$\nu_n'^2/2 > \operatorname*{Min}_{\dim(\mathfrak{N}+\mathfrak{P})\leqq j+k-2} \left[\operatorname*{Max}_{f\in\mathfrak{H}\ominus\mathfrak{N}} (D_1^2 f, f) + \operatorname*{Max}_{f\in\mathfrak{H}\ominus\mathfrak{P}} (D_2^2 f, f) \right]$$

which in turn implies

$$\nu_n'^2/2 > \operatorname*{Min}_{\dim(\mathfrak{N}+\mathfrak{P})\leqq j+k-2} \left[\operatorname*{Max}_{f\in\mathfrak{H}\ominus(\mathfrak{N}\oplus\mathfrak{P})} (D_1^2 f + D_2^2 f, f) \right],$$

f being restricted in every case to satisfy $|f| = 1$. But then, since $D_1^2 + D_2^2 - D^2/2$ is nonnegative definite, we have

$$\nu_n'^2 > \operatorname*{Min}_{\dim(\mathfrak{M})\leqq j+k-2} \operatorname*{Max}_{\substack{f\in\mathfrak{H}\ominus\mathfrak{M}\\ |f|=1}} (D^2 f, f)$$

which contradicts (1) unless $j + k - 2$ is greater than $n - 1$, or unless $j + k$ exceeds $n + 1$. Thus (4) holds for $j + k \leqq n + 1$, as stated, and we have

$$\nu_n' \leqq 2(\lambda_j' + \mu_k') \qquad \text{for } j + k \leqq n + 1. \tag{5}$$

Now let r_n be $n/2$ if n is even, $(n + 1)/2$ if n is odd. Then, from (5) we have

$$\nu_n' \leqq 2(\lambda_{r_n}' + \mu_{r_n}') \tag{6}$$

Moreover, the sequences $\{\lambda_n''\}$, $\{\mu_n''\}$ defined by the equations

$$\lambda_{2k-1}'' = 0, \quad \lambda_{2k}'' = \lambda_k', \qquad k = 1, 2, \cdots,$$

$$\mu_{2k-1}'' = 0, \quad \mu_{2k}'' = \mu_k', \qquad k = 1, 2, \cdots,$$

belong to $\mathfrak{F}$ by Lemma 1.2; and since

$$\lambda_{r_n}' = \lambda_n'' + \lambda_{\pi(n)}'', \qquad \mu_{r_n}' = \mu_n'' + \pi_{\pi(n)}'', \qquad n = 1, 2, \cdots$$

where $\pi(2k-1) = 2k$, $\pi(2k) = 2k-1$, it follows from conditions (i) and (iii) of Definition 1.1 that the sequences $\{\lambda'_{r_n}\}$ and $\{\mu'_{r_n}\}$ both belong to $\mathfrak{J}$. But then, by (6) and conditions (ii) and (iii) of our definition, $\{\nu'_n\}$ belongs to $\mathfrak{J}$, and this completes the proof of our assertion that $A + B$ belongs to $\mathcal{I}$.

Now let A be an arbitrary element of $\mathcal{I}$, $D_1 = (A^*A)^{\frac{1}{2}}$, X an arbitrary element of $\mathfrak{B}$, $D_2^2 = (A^*X^*XA)^{\frac{1}{2}}$. Since A is in $\mathcal{T}$, XA and D_2 are also. Let $\{\lambda_n\}$, $\{\mu_n\}$, $\{\lambda'_n\}$, $\{\mu'_n\}$ have the same meanings as above with reference to D_1 and D_2. Then $\{\lambda'_n\}$ is in $\mathfrak{J}$ and we can show that XA is in $\mathcal{I}$ by showing that $\{\mu'_n\}$ is in $\mathfrak{J}$; the arguments here are the same as above.

To establish the latter, we again apply the theorem of Courant;

$$\lambda_n'^2 = \operatorname*{Min}_{\dim(\mathfrak{M}) \leqq n-1} \operatorname*{Max}_{\substack{f \in \mathfrak{H} \ominus \mathfrak{M} \\ |f|=1}} (D_1^2 f, f),$$

$$\mu_n'^2 = \operatorname*{Min}_{\dim(\mathfrak{M}) \leqq n-1} \operatorname*{Max}_{\substack{f \in \mathfrak{H} \ominus \mathfrak{M} \\ |f|=1}} (D_2^2 f, f).$$

But $(D_1^2 f, f) = (Af, Af)$, $(D_2^2 f, f) = (XAf, XAf)$ and (XAf, XAf) is bounded by $N^2(Af, Af)$ for some integer N. Thus we have

$$\mu'_n \leqq N\lambda'_n,$$

and as $\{N\lambda'_n\}$ is clearly in $\mathfrak{J}$, so also is $\{\mu'_n\}$. Hence XA belongs to $\mathcal{I}$.

Thus $\mathcal{I}$ is a left ideal and in view of Theorem 1.2, we can show that $\mathcal{I}$ is two-sided by showing that $\mathcal{I}$ is closed with respect to the operation *. But this is an immediate consequence of Lemmas 1.1 and 1.2 and the well-known fact that $(A^*A)^{\frac{1}{2}}$ and $(AA^*)^{\frac{1}{2}}$ have the same positive characteristic values, each with the same multiplicity.[15]

The concluding assertion of the theorem is obvious.

THEOREM 1.7. *Let $\mathcal{F}$ denote the class of all operators A in $\mathfrak{B}$ such that $\mathfrak{R}(A)$ has a finite dimension number. Then $\mathcal{F}$ is a two-sided ideal in $\mathfrak{B}$. If $\mathcal{I}$ is an arbitrary two-sided ideal, in $\mathfrak{B}$, then $\mathcal{I} = (0)$ or $\mathcal{I} \supseteq \mathcal{F}$.*

To establish the first assertion, we have only to note that the class $\mathfrak{F}$ of all sequences in $\mathfrak{T}$ with only a finite number of terms different from zero constitutes an ideal set. To establish the second we first observe that any ideal set different from the one containing only the sequence all of whose terms are zero, contains a sequence $\{\lambda_n\}$ with $\lambda_1 \neq 0$, $\lambda_n = 0$ for $n \neq 1$, by virtue of (i) and(iii). Hence, by (ii) and (iii) it contains all sequences of this sort and hence, by (i) and (ii), contains $\mathfrak{F}$.

It is worthwhile to observe here that the effect of Theorem 1.6 is to establish an isomorphism between the lattice L_1 of two-sided ideals in $\mathfrak{B}$ and a certain sub-lattice L_2 of the lattice of ideals in the ring $\mathfrak{B}$ of all bounded sequences of complex numbers. To make this clear we require two preliminary results concerning $\mathfrak{B}$; the first theorem is the analogue for $\mathfrak{B}$ of Theorem 1.5 for $\mathfrak{B}$ the second is analogous to Theorems 1.3 and 1.4.

[15] By [12], Satz 7, for example.

sequence $\{\nu_n'^2\}$ associated with the operator D^2 in terms of maxima and minima. For our purposes, we may state this result as follows:

$$\nu_n'^2 = \operatorname*{Min}_{\dim(\mathfrak{M})\leqq n-1} \operatorname*{Max}_{\substack{f\in\mathfrak{H}\ominus\mathfrak{M}\\ |f|=1}} (D^2 f, f) \tag{1}$$

where $\dim(\mathfrak{M})$ is the dimension number of $\mathfrak{M}$. Similarly,

$$\lambda_n'^2 = \operatorname*{Min}_{\dim(\mathfrak{N})\leqq n-1} \operatorname*{Max}_{\substack{f\in\mathfrak{H}\ominus\mathfrak{N}\\ |f|=1}} (D_1^2 f, f), \tag{2}$$

$$\mu_n'^2 = \operatorname*{Min}_{\dim(\mathfrak{P})\leqq n-1} \operatorname*{Max}_{\substack{f\in\mathfrak{H}\ominus\mathfrak{P}\\ |f|=1}} (D_2^2 f, f). \tag{3}$$

On the basis of these relations we are able to conclude that the inequality

$$\nu_n'^2 \leqq 2(\lambda_j'^2 + \mu_k'^2) \tag{4}$$

is valid provided we have $j + k \leqq n + 1$. For if $\nu_n'^2 > 2(\lambda_j'^2 + \mu_k'^2)$ holds, we have, in view of (2) and (3),

$$\nu_n'^2/2 > \operatorname*{Min}_{\dim(\mathfrak{N})\leqq j-1} \operatorname*{Max}_{f\in\mathfrak{H}\ominus\mathfrak{N}} (D_1^2 f, f) + \operatorname*{Min}_{\dim(\mathfrak{P})\leqq k-1} \operatorname*{Max}_{f\in\mathfrak{H}\ominus\mathfrak{P}} (D_2^2 f, f)$$

and this implies

$$\nu_n'^2/2 > \operatorname*{Min}_{\dim(\mathfrak{N}+\mathfrak{P})\leqq j+k-2} \left[\operatorname*{Max}_{f\in\mathfrak{H}\ominus\mathfrak{N}} (D_1^2 f, f) + \operatorname*{Max}_{f\in\mathfrak{H}\ominus\mathfrak{P}} (D_2^2 f, f)\right]$$

which in turn implies

$$\nu_n'^2/2 > \operatorname*{Min}_{\dim(\mathfrak{N}+\mathfrak{P})\leqq j+k-2} \left[\operatorname*{Max}_{f\in\mathfrak{H}\ominus(\mathfrak{N}\oplus\mathfrak{P})} (D_1^2 f + D_2^2 f, f)\right],$$

f being restricted in every case to satisfy $|f| = 1$. But then, since $D_1^2 + D_2^2 - D^2/2$ is nonnegative definite, we have

$$\nu_n'^2 > \operatorname*{Min}_{\dim(\mathfrak{M})\leqq j+k-2} \operatorname*{Max}_{\substack{f\in\mathfrak{H}\ominus\mathfrak{M}\\ |f|=1}} (D^2 f, f)$$

which contradicts (1) unless $j + k - 2$ is greater than $n - 1$, or unless $j + k$ exceeds $n + 1$. Thus (4) holds for $j + k \leqq n + 1$, as stated, and we have

$$\nu_n' \leqq 2(\lambda_j' + \mu_k') \qquad \text{for } j + k \leqq n + 1. \tag{5}$$

Now let r_n be $n/2$ if n is even, $(n + 1)/2$ if n is odd. Then, from (5) we have

$$\nu_n' \leqq 2(\lambda_{r_n}' + \mu_{r_n}') \tag{6}$$

Moreover, the sequences $\{\lambda_n''\}$, $\{\mu_n''\}$ defined by the equations

$$\lambda_{2k-1}'' = 0, \qquad \lambda_{2k}'' = \lambda_k', \qquad k = 1, 2, \cdots,$$

$$\mu_{2k-1}'' = 0, \qquad \mu_{2k}'' = \mu_k', \qquad k = 1, 2, \cdots,$$

belong to $\mathfrak{Z}$ by Lemma 1.2; and since

$$\lambda_{r_n}' = \lambda_n'' + \lambda_{\pi(n)}'', \qquad \mu_{r_n} = \mu_n'' + \pi_{\pi(n)}'', \qquad n = 1, 2, \cdots$$

where $\pi(2k-1) = 2k$, $\pi(2k) = 2k-1$, it follows from conditions (i) and (iii) of Definition 1.1 that the sequences $\{\lambda'_{r_n}\}$ and $\{\mu'_{r_n}\}$ both belong to $\mathfrak{J}$. But then, by (6) and conditions (ii) and (iii) of our definition, $\{\nu'_n\}$ belongs to $\mathfrak{J}$, and this completes the proof of our assertion that $A + B$ belongs to $\mathfrak{I}$.

Now let A be an arbitrary element of $\mathfrak{I}$, $D_1 = (A^*A)^{\frac{1}{2}}$, X an arbitrary element of $\mathfrak{B}$, $D_2^2 = (A^*X^*XA)^{\frac{1}{2}}$. Since A is in $\mathfrak{T}$, XA and D_2 are also. Let $\{\lambda_n\}$, $\{\mu_n\}$, $\{\lambda'_n\}$, $\{\mu'_n\}$ have the same meanings as above with reference to D_1 and D_2. Then $\{\lambda'_n\}$ is in $\mathfrak{J}$ and we can show that XA is in $\mathfrak{I}$ by showing that $\{\mu'_n\}$ is in $\mathfrak{J}$; the arguments here are the same as above.

To establish the latter, we again apply the theorem of Courant;

$$\lambda_n'^2 = \operatorname*{Min}_{\dim(\mathfrak{M}) \leqq n-1} \operatorname*{Max}_{\substack{f \in \mathfrak{H} \ominus \mathfrak{M} \\ |f|=1}} (D_1^2 f, f),$$

$$\mu_n'^2 = \operatorname*{Min}_{\dim(\mathfrak{M}) \leqq n-1} \operatorname*{Max}_{\substack{f \in \mathfrak{H} \ominus \mathfrak{M} \\ |f|=1}} (D_2^2 f, f).$$

But $(D_1^2 f, f) = (Af, Af)$, $(D_2^2 f, f) = (XAf, XAf)$ and (XAf, XAf) is bounded by $N^2(Af, Af)$ for some integer N. Thus we have

$$\mu'_n \leqq N\lambda'_n,$$

and as $\{N\lambda'_n\}$ is clearly in $\mathfrak{J}$, so also is $\{\mu'_n\}$. Hence XA belongs to $\mathfrak{I}$.

Thus $\mathfrak{I}$ is a left ideal and in view of Theorem 1.2, we can show that $\mathfrak{I}$ is two-sided by showing that $\mathfrak{I}$ is closed with respect to the operation *. But this is an immediate consequence of Lemmas 1.1 and 1.2 and the well-known fact that $(A^*A)^{\frac{1}{2}}$ and $(AA^*)^{\frac{1}{2}}$ have the same positive characteristic values, each with the same multiplicity.[15]

The concluding assertion of the theorem is obvious.

THEOREM 1.7. *Let $\mathfrak{F}$ denote the class of all operators A in $\mathfrak{B}$ such that $\mathfrak{R}(A)$ has a finite dimension number. Then $\mathfrak{F}$ is a two-sided ideal in $\mathfrak{B}$. If $\mathfrak{I}$ is an arbitrary two-sided ideal, in $\mathfrak{B}$, then $\mathfrak{I} = (0)$ or $\mathfrak{I} \supseteq \mathfrak{F}$.*

To establish the first assertion, we have only to note that the class $\mathfrak{F}$ of all sequences in $\mathfrak{T}$ with only a finite number of terms different from zero constitutes an ideal set. To establish the second we first observe that any ideal set different from the one containing only the sequence all of whose terms are zero, contains a sequence $\{\lambda_n\}$ with $\lambda_1 \neq 0$, $\lambda_n = 0$ for $n \neq 1$, by virtue of (i) and(iii). Hence, by (ii) and (iii) it contains all sequences of this sort and hence, by (i) and (ii), contains $\mathfrak{F}$.

It is worthwhile to observe here that the effect of Theorem 1.6 is to establish an isomorphism between the lattice L_1 of two-sided ideals in $\mathfrak{B}$ and a certain sub-lattice L_2 of the lattice of ideals in the ring $\mathfrak{B}$ of all bounded sequences of complex numbers. To make this clear we require two preliminary results concerning $\mathfrak{B}$; the first theorem is the analogue for $\mathfrak{B}$ of Theorem 1.5 for $\mathfrak{B}$ the second is analogous to Theorems 1.3 and 1.4.

[15] By [12], Satz 7, for example.

THEOREM 1.8. *Let $\mathfrak{B}$ be the ring of all bounded sequences of complex numbers, $\mathfrak{J}$ an ideal in $\mathfrak{B}$. Let $\mathfrak{J}_0$ be the set of all sequences $\{|\lambda_n|\}$ such that $\{\lambda_n\}$ is in $\mathfrak{J}$. Then $\mathfrak{J}$ contains $\mathfrak{J}_0$, and if $\mathfrak{J}_1$ is an ideal in $\mathfrak{B}$ containing $\mathfrak{J}_0$, then $\mathfrak{J}_1 \supseteq \mathfrak{J}$.*

The proof is straightforward and is left to the reader.

THEOREM 1.9. *Let L_2 be the lattice of all ideals $\mathfrak{J}$ in $\mathfrak{B}$ which satisfy the following condition; if $\{a_n\}$ is in $\mathfrak{J}$, and π is a permutation of the positive integers $\{a_{\pi(n)}\}$ is in $\mathfrak{J}$. Then the set $\mathfrak{T}$ of sequences convergent to zero belongs to L_2, and if $\mathfrak{J}$ is an arbitrary element of L_2, either $\mathfrak{J} = \mathfrak{B}$ or $\mathfrak{J} \subseteq \mathfrak{T}$.*

The first assertion is obvious. Now if $\mathfrak{J}$ is a member of L_2 which is not contained in $\mathfrak{T}$, $\mathfrak{J}$ contains a sequence $\{a_n\}$ which has a subsequence $\{a_{n_k}\}$ such that $\{a_{n_k}^{-1}\}$ is bounded, and which converges to a number different from zero. Furthermore, the sequence $\{b_n\}$ with $b_{n_k} = a_{n_k}$, $b_n = 0$, $n \neq n_k$ can be written $\{b_n\} = \{c_n\}\{a_n\}$, where $c_{n_k} = 1$, $c_n = 0$, $n \neq n_k$, and hence $\{b_n\}$ belongs to $\mathfrak{J}$. We now write the sequence $\{b_n\}$ in the form $\{b_n\} = \{\rho_n e^{i\theta_n}\}$, $\rho_n > 0$, and observe that $\{\rho_n\}$ also belongs to $\mathfrak{J}$ by Theorem 1.8. But then by the same sort of argument that was used to prove Lemma 1.1, we can show that $\{\rho_{n_k}\}$ belongs to $\mathfrak{J}$, and hence $\{a_{n_k}\}$ does also. Hence $\{e_k\} = \{a_{n_k}\}\{a_{n_k}^{-1}\}$ belongs to $\mathfrak{J}$ and $e_k = 1$, $k = 1, 2, \cdots$. Thus $\mathfrak{J} = \mathfrak{B}$.

THEOREM 1.10. *Let $\mathfrak{J}$ be a member of L_2, $\mathfrak{J}_0$ the subset of $\mathfrak{J}$ defined in Theorem 1.8. Then either $\mathfrak{J}_0 = \mathfrak{B}_0$, where $\mathfrak{B}_0$ the set of all sequences of nonnegative numbers in $\mathfrak{B}$, or every sequence in $\mathfrak{J}_0$ is convergent to zero and $\mathfrak{J}_0$ is an ideal set in the sense of Definition* 1.1.

THEOREM 1.10 follows at once from Theorems 1.8 and 1.9.

THEOREM 1.11. *Let L_1 be the lattice of all two-sided ideals in $\mathfrak{B}$, let the set $\mathfrak{B}_0$ be called the spectral set of $\mathfrak{B}$, and let the class of ideal sets be extended to include $\mathfrak{B}_0$. Then L_1 is lattice-isomorphic to the extended class of ideal sets, each member of L_1 corresponding under this isomorphism to its spectral set. Similarly, the lattice L_2 is lattice-isomorphic to the extended class of all ideal sets, each member $\mathfrak{J}$ of L_2 corresponding to its $\mathfrak{J}_0$. Thus L_1 and L_2 are lattice-isomorphic and under this isomorphism $\mathfrak{B}$ corresponds to $\mathfrak{B}$, $\mathfrak{T}$ to the class $\mathfrak{T}$ of all sequences convergent to zero, $\mathfrak{F}$ to the class $\mathfrak{F}$ of all sequences with only a finite number of terms different from zero.*

Theorem 1.11 is obvious on the basis of preceding results and we omit the proof. Before proceeding, however, we wish to make the following observations: The ring $\mathfrak{B}$ can be imbedded in $\mathfrak{B}$ in a very simple way; we have merely to choose a complete orthonormal set $\{\varphi_n\}$ in $\mathfrak{H}$ and identify the element $\{a_n\}$ of $\mathfrak{B}$ with the closed linear operator A in $\mathfrak{H}$ which is defined on $\{\varphi_n\}$ by the equations $A\varphi_n = a_n\varphi_n$, $n = 1, 2, \cdots$. Moreover, if we consider $\mathfrak{B}$ in terms of this identification, each two-sided ideal $\mathfrak{I}$ in $\mathfrak{B}$ corresponds under the isomorphism of Theorem 1.11 merely to its intersection with $\mathfrak{B}$. This suggests an alternative attack on the problem solved by Theorem 1.6; however, in so far as we can determine, it is not possible to devise any essentially different proof of that theorem on this basis.

2. Congruences in $\mathfrak{B}$

We pass now to the study of congruences modulo an ideal $\mathfrak{I}$ in $\mathfrak{B}$. Following the standard procedure of abstract algebra, we consider the class $\mathfrak{B}/\mathfrak{I}$ whose elements $\alpha, \beta, \cdots$ are the residue classes of $\mathfrak{B}$ with respect to $\mathfrak{I}$; by definition, two members A and B of $\mathfrak{B}$ belong to the same element α of $\mathfrak{B}/\mathfrak{I}$ if and only if $A - B$ is in $\mathfrak{I}$. If α and β are arbitrary elements of $\mathfrak{B}/\mathfrak{I}$, we define $\alpha + \beta$ as the class of all elements $A + B$ of $\mathfrak{B}$ such that A is in α, B in β; similarly, we define $\alpha\beta$ as the class of all AB in $\mathfrak{B}$ such that A is in α, B in β. Then, from the general theorem[16] which is controlling in such situations, we have

THEOREM 2.1. *If α and β are elements of $\mathfrak{B}/\mathfrak{I}$, so also are $\alpha + \beta$ and $\alpha\beta$. With addition and multiplication defined in this way $\mathfrak{B}/\mathfrak{I}$ is a ring; that is to say, $\mathfrak{B}/\mathfrak{I}$ is a commutative group with respect to the operation $+$, and further, the following formal laws are satisfied*:

$$(\alpha\beta)\gamma = \alpha(\beta\gamma), \qquad \alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma, \qquad (\beta + \gamma)\alpha = \beta\alpha + \gamma\alpha.$$

Moreover, $\mathfrak{B}/\mathfrak{I}$ possesses a unit.

It may be noted that except in the case $\mathfrak{I} = \mathfrak{B}$, the subclass of $\mathfrak{B}/\mathfrak{I}$ each of whose elements contains a scalar multiple of the identity in $\mathfrak{B}$, is isomorphic to the class of scalar multiples of the identify in $\mathfrak{B}$, since no two of these elements of $\mathfrak{B}$ can have difference in $\mathfrak{I}$ unless they are identical. It is convenient therefore to use italic letters for these elements as well as for the corresponding elements of $\mathfrak{B}$. In addition we shall use the symbol 1 for the unit in $\mathfrak{B}/\mathfrak{I}$; that is, for the element of $\mathfrak{B}/\mathfrak{I}$ whose members have the form $I + T$, where I is the identity in $\mathfrak{B}$ and T belongs to $\mathfrak{I}$.

It is worth pointing out here that the ring $\mathfrak{B}/\mathfrak{I}$, $\mathfrak{I} \neq \mathfrak{B}$, is certainly non-commutative; to verify this one needs only to consider two orthogonal projections E and F whose ranges are Hilbert spaces and whose sum is the identity, a partially isometric operator W which maps E on F, and the operator $W^*W - WW^* = E - F$ which belongs to no ideal except $\mathfrak{B}$ itself. Later we shall show that the center of $\mathfrak{B}/\mathfrak{I}$, $\mathfrak{I} \neq \mathfrak{B}$, is the set of all scalar multiples of unity (Theorem 2.9).

THEOREM 2.2. *If α is an arbitrary element of $\mathfrak{B}/\mathfrak{I}$, the class α^* of all members A^* of $\mathfrak{B}$ such that A is in α is in $\mathfrak{B}/\mathfrak{I}$ also. The operation * so defined in $\mathfrak{B}/\mathfrak{I}$ obeys the following laws*:

$$\alpha^{**} = \alpha, \qquad (\alpha + \beta)^* = \alpha^* + \beta^*, \qquad (\alpha\beta)^* = \beta^*\alpha^*.$$

That α^* is in $\mathfrak{B}/\mathfrak{I}$ follows at once from the fact that $\mathfrak{I}^* = \mathfrak{I}$; and the three laws stated in the theorem are readily verified on the basis of their validity in $\mathfrak{B}$.

Thus we see that the rings $\mathfrak{B}/\mathfrak{I}$ have all of the formal properties of matrix algebras and are homomorphs of $\mathfrak{B}$ with respect to the operations $+$, $\cdot$, *, $a\cdot$;

[16] [1], pp. 252–253. The missing details necessary for our purposes are readily supplied. Cf. the discussion of the commutative case in [19].

moreover, they are of course the only homomorphs of $\mathfrak{B}$ with respect to these operations.

It is now desirable to consider these homomorphisms with respect to the following important notions in operator theory; for an operator to be self-adjoint, to be idempotent, to be partially isometric, to be unitary. Hence we are led to define these concepts in $\mathfrak{B}/\mathfrak{I}$ without explicit reference to their meanings in $\mathfrak{B}$.

DEFINITION 2.1. *An element α of $\mathfrak{B}/\mathfrak{I}$ is called self-adjoint if $\alpha = \alpha^*$; a self-adjoint element ϵ of $\mathfrak{B}/\mathfrak{I}$ is called idempotent if $\epsilon^2 = \epsilon$; an element ω of $\mathfrak{B}/\mathfrak{I}$ is called partially isometric if $\omega^*\omega = \epsilon$ is idempotent, unitary if $\omega^*\omega = \omega\omega^* = 1$.*

It is easy to see that under the homomorphisms $\mathfrak{B} \to \mathfrak{B}/\mathfrak{I}$ the image of every-self adjoint operator is self-adjoint and that analogous assertions hold for projections, partially isometric operators, and unitary operators. We shall now show that with reference to the first two of these concepts the converse statements are also true.

THEOREM 2.3. *If α is a self-adjoint element of $\mathfrak{B}/\mathfrak{I}$, α contains a self-adjoint member of $\mathfrak{B}$, and conversely.*

Let α be self-adjoint, A an element of α. Then $A^* - A$ is in $\mathfrak{I}$ and hence $A + (A^* - A)/2 = (A + A^*)/2$ is in α. The converse, as we have already noted, is obvious.

THEOREM 2.4. *If ϵ is an idempotent element of $\mathfrak{B}/\mathfrak{I}$, there exists a projection E in $\mathfrak{B}$ which belongs to ϵ, and conversely.*

The theorem is obvious for $\mathfrak{I} = \mathfrak{B}$; we assume therefore $\mathfrak{I} \subseteq \mathfrak{T}$. By Theorem 2.3, ϵ contains a self-adjoint transformation A, and since ϵ is idempotent, $A^2 - A = A(A - I)$ is in $\mathfrak{I}$ and thus in $\mathfrak{T}$. Hence $A^2 - A$ can be reduced to diagonal form, and therefore A can also; $A\varphi_n = \lambda_n\varphi_n$, $n = 1, 2, \cdots$, where $\{\varphi_n\}$ is a complete orthonormal set in $\mathfrak{H}$. But then it follows that $\{\lambda_n\}$ contains a subsequence $\{\lambda_n^{(0)}\}$ convergent to zero, and such that the remaining terms of $\{\lambda_n\}$ form a subsequence, say $\{\lambda_n^{(1)}\}$, convergent to 1, since under any other circumstances $A(A - I)$ would fail to be totally continuous. Moreover, we can clearly assume that $\{\lambda_n^{(1)}\}$ contains no terms with the value zero and that $\{\lambda_n^{(0)}\}$ contains no terms with the value unity.

Now let $\mathfrak{M}_0$ be the subspace of $\mathfrak{H}$ determined by the characteristic elements of A corresponding to terms of $\{\lambda_n^{(0)}\}$, $\mathfrak{M}_1$ the subspace determined by the other characteristic elements of A, E_0 and E_1 the projections with ranges $\mathfrak{M}_0$ and $\mathfrak{M}_1$, respectively. We shall show that E_1 belongs to ϵ. To do this we note first that in $\mathfrak{M}_0$, $A - I$ induces a transformation with bounded inverse. Hence if B is equal to this inverse in $\mathfrak{M}_0$ and to zero in $\mathfrak{M}_1$,

$$E_0(A^2 - A)BE_0 = E_0AE_0$$

is in $\mathfrak{I}$, since $A^2 - A$ is. Similarly, it follows that

$$E_1(A - I)E_1 = E_1AE_1 - E_1$$

is in $\mathfrak{I}$. But then, adding the right members of the two preceding equations we find that

$$E_0AE_0 + E_1AE_1 - E_1 = A - E_1$$

is in $\mathcal{I}$, and from this it follows that E_1 belongs to ϵ.

Again the converse part of the theorem is obvious, so the proof is complete.

THEOREM 2.5. *Let $\mathcal{I}$ be an ideal in $\mathcal{B}$, $\mathcal{I}_0$ the set of all nonnegative definite self-adjoint elements of $\mathcal{I}$, $\mathcal{I}_0^2$ the class of all squares of elements of $\mathcal{I}_0$. Then a necessary and sufficient condition that every partially isometric element ω of $\mathcal{B}/\mathcal{I}$ contain a partially isometric transformation W is that $\mathcal{I}_0 = \mathcal{I}_0^2$.*

Again the case $\mathcal{I} = \mathcal{B}$ is trivial, so we assume $\mathcal{I} \subseteq \mathcal{T}$. Let ω be partially isometric, V an element of ω, $V = UB$ its canonical decomposition. Then $V^*V = B^2$ belongs to an idempotent element ϵ of $\mathcal{B}/\mathcal{I}$. Hence, if we identify B^2 with the self-adjoint transformation A which appears in the proof of Theorem 2.4, we can invoke that theorem to establish the existence of a projection E such that $B^2 - E$ is in $\mathcal{I}$. Moreover, an inspection of the proof reveals also that B^2 commutes with E and induces in the range of E a transformation with bounded inverse. In particular, this implies that E has for its range a subspace of the initial set of U and thus that UE is partially isometric, since $EU^*UE = E$.[17] We shall now show that under the condition of the theorem $B - E$ is in $\mathcal{I}$. We note first that since B^2 and E commute, we have $B^2 - E = (B - E)(B + E)$, and since B is nonnegative, $B + E$ induces in $\mathfrak{R}(E)$ a transformation with bounded inverse. Hence, if C is equal to this inverse in $\mathfrak{R}(E)$, and to zero in $\mathfrak{H} \ominus \mathfrak{R}(E)$, we have $(B^2 - E)C = EB - E$, and $EB - E$ is in $\mathcal{I}$. Moreover $(I - E)(B^2 - E) = (I - E)B^2$ is in $\mathcal{I}$. But if $\mathcal{I}$ has the property described in the theorem, $[(I - E)B^2]^\dagger = (I - E)B$ is in $\mathcal{I}$, and thus

$$(I - E)B + EB - E = B - E$$

is in $\mathcal{I}$ as we wished to show. But then $V - UE = U(B - E)$ is in $\mathcal{I}$ and $W = UE$, which we have already shown to be partially isometric, belongs to ω.

It remains therefore for the converse part of the theorem to be proved. To this end, we suppose that $\mathcal{I}$ contains a nonnegative definite self-adjoint transformation B^2 such that B does not belong to $\mathcal{I}$, and denote by ω the congruence class in $\mathcal{B}/\mathcal{I}$ to which B belongs. Then, since $B^*B = B^2$ belongs to $\mathcal{I}$, we have ω partially isometric by definition. But if V is a partially isometric transformation in ω, $V^*V = E$ must be congruent to B^2 modulo $\mathcal{I}$, which is to say that E is congruent to zero modulo $\mathcal{I}$. But this implies that E has range with finite dimension number and the range of E is the initial set of V. Thus V is in $\mathcal{F}$ and hence in $\mathcal{I}$. However, $V - B$ is in $\mathcal{I}$ since V is in ω, and as B is by assumption not in $\mathcal{I}$, we have a contradiction. Hence, the condition of the theorem is necessary as well as sufficient.

We may note in passing that the ideals (0), $\mathcal{F}$, $\mathcal{T}$, $\mathcal{B}$ all satisfy the condition of Theorem 2.5, but that these are not the only ideals which do so. Consider, for example, the class of all sequences $\{\lambda_n\}$ of nonnegative numbers such that

[17] By [9], Lemma 4.3.2.

$\sum_{n=1}^{\infty} \lambda_n^p$ converges for some p. It is easily seen that this class is an ideal set in the sense of Definition 1.1. and that the corresponding ideal in B has the property of Theorem 2.5.

For the sake of completeness, we state the following obvious theorem:

THEOREM 2.6. *If W is a partially isometric member of $\mathfrak{B}$, the congruence class of W in $\mathfrak{B}/\mathfrak{J}$ is partially isometric in $\mathfrak{B}/\mathfrak{J}$.*

THEOREM 2.7. *Let $\mathfrak{J}$ be an ideal in $\mathfrak{B}$. Then, if ω is a unitary element of $\mathfrak{B}/\mathfrak{J}$, ω contains a maximal partially isometric transformation[18] with deficiency-index $(0, n)$ or $(n, 0)$, $n < \aleph_0$.*

Again the case $\mathfrak{J} = \mathfrak{B}$ is trivial, so we assume $\mathfrak{J} \neq \mathfrak{B}$. Let ω be unitary, V a member of ω, $V = UB$ its canonical decomposition. Since ω is also partially isometric, the first part of the proof of Theorem 2.5 applies to yield the following results: there exists a projection E with range in the initial set of U, and which commutes with B, such that $V^*V - E = B^2 - E$ and $EB - E$ are in $\mathfrak{J}$. But since V^*V is congruent to I modulo $\mathfrak{J}$, $I - E$ is in $\mathfrak{J}$ and thus $(I - E)B$ is in $\mathfrak{J}$. Therefore

$$B - E = (I - E)B + EB - E$$

is in $\mathfrak{J}$. But then $V - UE = UB - UE$ is in $\mathfrak{J}$. Hence $W = UE$ is a partially isometric operator which belongs to ω. Moreover, since ω is unitary, $I - WW^*$ is in $\mathfrak{J}$ and hence, since this operator is a projection and belongs to $\mathfrak{J}$, its range must have a finite dimension number. Similarly, $I - W^*W$ has range with a finite dimension number. Thus both the initial and final sets of W have orthogonal complements with finite dimension numbers. Therefore, if W_1 is the contraction of W with domain $\mathfrak{R}(E)$ and X a maximal partially isometric extension of W_1, X has the property required in the theorem and $X - W$ is in $\mathfrak{J}$. Thus X belongs to ω, and the theorem is proved.

It is important to observe that every unitary element of $\mathfrak{B}/\mathfrak{J}$ does not contain a unitary member of $\mathfrak{B}$, except in the trivial cases $\mathfrak{J} = \mathfrak{B}$, $\mathfrak{J} = (0)$.

To prove this, we consider an isometric transformation X with deficiency-index $(0, n)$, $n < \aleph_0$, and the congruence class ω modulo $\mathfrak{J}$, to which X belongs. Then ω is clearly unitary in $\mathfrak{B}/\mathfrak{J}$ provided $\mathfrak{J} \neq (0)$. Now suppose U is a unitary transformation in ω. Then $U - X$ is in $\mathfrak{F}$, and thus in $\mathfrak{T}$, if $\mathfrak{J} \neq \mathfrak{B}$. Hence $I - U^{-1}X$ is in $\mathfrak{J}$ and $U^{-1}X$ also has deficiency-index $(0, n)$. But by a lemma which the author has proved elsewhere, this is possible if and only if $n = 0$.[19]

THEOREM 2.8. *Let $\mathfrak{J}$ be an ideal in $\mathfrak{B}$ different from (0), W a partially isometric operator in $\mathfrak{B}$ with deficiency-index (m, n) $m, n < \aleph_0$. Then the congruence class ω in $\mathfrak{B}/\mathfrak{J}$ to which W belongs is unitary.*

If W has the properties stated then $I - W^*W$ and $I - WW^*$ are projections

[18] We call a partially isometric operator maximal if the isometric transformation which determines it is maximal. Similarly, we shall have occasion to refer to the deficiency-index of a partially isometric operator.

[19] [3], Lemma 4.1.

which belong to $\mathcal{F}$. Thus, since $\mathcal{F} \subseteq \mathcal{I}$ by Theorem 1.7, both of these operators belong to $\mathcal{I}$ and ω is unitary by definition.

We conclude this section with

THEOREM 2.9. *Let $\mathcal{I}$ be an ideal in $\mathcal{B}$, $\mathcal{I} \neq \mathcal{B}$. Then the center of $\mathcal{B}/\mathcal{I}$, that is, the set of all elements of $\mathcal{B}/\mathcal{I}$ which commute with every element of $\mathcal{B}/\mathcal{I}$, is the set of all elements $\lambda \cdot 1$, where λ is a complex number.*

It is clear that the center contains the set of all scalar multiples of the identity; hence we need only show that it contains no other members.

We begin by showing that it is sufficient to consider merely the self-adjoint members of the center. For suppose α belongs to the center. Then $\alpha\beta^* - \beta^*\alpha = 0$ for all β in $\mathcal{B}/\mathcal{I}$ and hence $(\alpha\beta^* - \beta^*\alpha)^* = \beta\alpha^* - \alpha^*\beta = 0$ for all β in $\mathcal{B}/\mathcal{I}$. Thus α^* belongs to the center, and consequently the self-adjoint elements $\alpha + \alpha^*$ and $i(\alpha - \alpha^*)$ do also. Now suppose $\alpha + \alpha^* = \lambda \cdot 1$, $i(\alpha - \alpha^*) = \mu \cdot 1$. Then, eliminating α^*, we have $\alpha = (\mu + i\lambda)/2i$. Hence we have only to prove that every self-adjoint member of the center is a scalar multiple of the identity.

In terms of operators in $\mathcal{B}$, this problem reduces to the following: to show that every self-adjoint operator A in $\mathcal{B}$ such that $AB - BA$ is in $\mathcal{I}$ for all B in $\mathcal{B}$ is of the form $T + \lambda I$, where T is in $\mathcal{I}$.

We consider first the case $\mathcal{I} = \mathcal{T}$. So we consider a self-adjoint operator A so that $AB - BA$ is totally continuous for all B in $\mathcal{B}$. If A is not of the form $T + \lambda I$, T in $\mathcal{T}$, the spectrum of A must contain two distinct points, each of which is either a limit point of the spectrum of A or a characteristic value of infinite multiplicity; for, otherwise, the spectrum of A consists solely of isolated characteristic values of finite multiplicity together with one point μ which is either a limit point or a characteristic value of infinite multiplicity, and in this case $A - \mu I$ clearly belongs to $\mathcal{T}$. Hence if $E(\lambda)$ is the resolution of the identity of A, there exist numbers λ_0, λ_1, λ_2, λ_3 in the spectrum of A, $\lambda_0 < \lambda_1 < \lambda_2 < \lambda_3$ such that $E(\lambda_1) - E(\lambda_0)$ and $E(\lambda_3) - E(\lambda_2)$ have ranges $\mathfrak{M}_1$ and $\mathfrak{M}_2$, respectively, which are Hilbert spaces.

Next let us consider the partially isometric operator W with initial set $\mathfrak{M}_1$ and final set $\mathfrak{M}_2$. We then have, for f in $\mathfrak{M}_1$

$$|WAf - AWf| \geqq |AWf| - |WAf|,$$

and thus, since we also have

$$|AWf| \geqq \lambda_2 |Wf| = \lambda_2 |f|, \quad |WAf| = |Af| \leqq \lambda_1 |f|$$

we obtain

$$|WAf - AWf| \geqq (\lambda_2 - \lambda_1) f.$$

Hence $(WA - AW)$ induces on $\mathfrak{M}_1$ a transformation with bounded inverse and therefore $(WA - AW)\mathfrak{M}_1$ is a Hilbert space. Consequently, by a lemma previously referred to, $WA - AW$ is not in $\mathcal{T}$. Hence the assumption that A is not of the form $T + \lambda I$, T in $\mathcal{T}$, is untenable, and the theorem is established for $\mathcal{I} = \mathcal{T}$.

Consider now an arbitrary ideal $\mathcal{J}$, $\mathcal{F} \subset \mathcal{J} \subset \mathcal{T}$, $\mathcal{J} \neq \mathcal{F}$, $\mathcal{J} \neq \mathcal{T}$. As before, we consider a self-adjoint operator A such that $AB - BA$ is in $\mathcal{J}$ for all B in $\mathfrak{B}$. Since $\mathcal{J} \subset \mathcal{T}$, it follows from the preceding result that A is of the form $T + \lambda I$, T in $\mathcal{T}$. Thus $TB - BT$ is in $\mathcal{J}$ for all B in $\mathfrak{B}$. Now let $\{\varphi_n\}$ be a complete orthonormal set of characteristic elements of T; $T\varphi_n = \lambda_n\varphi_n$, $n = 1, 2, \cdots$, and let us suppose that T is not in $\mathcal{J}$. Moreover, let us assume that the sequence $\{\varphi_n\}$ is so arranged that λ_1 is different from zero, while the positive terms of $\{|\lambda_n|\}$ are in monotone order. Let $\{\mu_n\}$ be a sequence of positive numbers in the spectral set of $\mathcal{J}$ with $\mu_1 < |\lambda_1|$; further let $\{\lambda_{n_k}\}$ be an infinite subsequence of $\{\lambda_n\}$ such that $0 < |\lambda_{n_k}| \leqq \mu_k$, $k = 1, 2, \cdots$.[20] Finally, let E be the projection with range $\mathfrak{M}$ determined by the orthonormal set $\{\varphi_{n_k}\}$ and let W be a partially isometric transformation with initial set $\mathfrak{H}$ and final set $\mathfrak{M}$.

Then $WT - TW$ is in $\mathcal{J}$. Hence $WTW^* - TWW^* = WTW^* - TE$ is in $\mathcal{J}$. But then, since $\{\lambda_{n_k}\}$ is in the spectral set of $\mathcal{J}$ by choice of that subsequence, TE is in $\mathcal{J}$. Hence WTW^* is in $\mathcal{J}$, and therefore $W^*WTW^*W = T$ is in $\mathcal{J}$ too, which is a contradiction.

It remains to prove the theorem for the case $\mathcal{J} = \mathcal{F}$. Let us suppose that $BA - AB$ is in $\mathcal{F}$ for all B in $\mathfrak{B}$ and that A is not in $\mathcal{F}$. Then A is of the form $T + \lambda I$, where T is in $\mathcal{T}$ and not in $\mathcal{F}$. Hence there exists an infinite orthonormal set $\{\varphi_n\}$ in $\mathfrak{H}$ such that $T\varphi_n = \lambda_n\varphi_n$, $\lambda_n \neq 0$, $n = 1, 2, \cdots$, $\lambda_n \neq \lambda_m$ if $m \neq n$. Hence, if U is defined by the equations

$$U\varphi_{2n} = \varphi_{2n-1}, \qquad U\varphi_{2n-1} = \varphi_{2n}, \qquad n = 1, 2, \cdots,$$

in the closed linear manifold determined by $\{\varphi_n\}$, $U = I$ in $\mathfrak{H} \ominus \mathfrak{M}$, $UTU^{-1} - T$ is not in $\mathcal{F}$ and hence $UT - TU$ is not either. Therefore the assumption that A is not in $\mathcal{F}$ leads to a contradiction and the theorem is proved for $\mathcal{J} = \mathcal{F}$.

3. A metric in $\mathfrak{B}/\mathcal{T}$

We now confine our attention to the case $\mathcal{J} = \mathcal{T}$, beginning with the definition of a norm in $\mathfrak{B}/\mathcal{T}$.

Throughout the remainder of the paper we employ the notation $|A|$ for the bound of the operator A of $\mathfrak{B}$.

DEFINITION 3.1. *Let α be an arbitrary element of $\mathfrak{B}/\mathcal{T}$. We define $|\alpha|$, called the norm of α, by the equation*

$$|\alpha| = \text{g. l. b.}_{A \in \alpha} |A|.$$

THEOREM 3.1. *The norm $|\alpha|$ in $\mathfrak{B}/\mathcal{T}$ has the following properties:*

(1) $|\alpha| \geqq 0$, *the equality sign holding if and only if* $\alpha = 0$;

(2) $|\alpha + \beta| \leqq |\alpha| + |\beta|$;

(3) $|\alpha\beta| \leqq |\alpha|\,|\beta|$;

(4) $|a\alpha| = |a|\,|\alpha|$;

(5) $|\alpha^*| = |\alpha|$;

(6) $|1| = 1$.

[20] If no such subsequence exists, T is in $\mathcal{F}$ and hence in $\mathcal{J}$.

The validity of the laws (2) – (5) is an immediate consequence of the definition of $|\alpha|$ in terms of the norm $|A|$ in $\mathfrak{B}$, and the fact that the latter function has those properties. The same is true of the assertion $|\alpha| \geqq 0$. Moreover, in view of (2) and (3) we can conclude that the set of elements α of $\mathfrak{B}/\mathfrak{T}$ for which $|\alpha| = 0$ is a two-sided ideal in $\mathfrak{B}/\mathfrak{T}$. Hence, since $\mathfrak{T}$ is a prime ideal in $\mathfrak{B}$, we have either $\alpha = 0$ when and only when $|\alpha| = 0$, or $\alpha = 0$ for all α in $\mathfrak{B}/\mathfrak{T}$.[21] Thus to complete the proof it is necessary only to establish (6). We have then to show that

$$\text{g. l. b.}_{T \in \mathfrak{T}} |I + T| = 1.$$

We note first

$$|I + (T + T^*)/2| \leqq |(I + T)/2| + |(I + T^*)/2| = |I + T|;$$

hence we need only show g. l. b. $|I + T| = 1$ for $T \in \mathfrak{T}$, $T = T^*$. But, if T is so restricted, we can find, for any $\epsilon > 0$, a real number λ with $|\lambda| < \epsilon$ such that $T\varphi = \lambda\varphi$ for some $\varphi \neq 0$ in $\mathfrak{H}$, by virtue of the spectral properties of T. Thus $|(I + T)\varphi| / |\varphi| > (1 - \epsilon)$ and hence $|I + T|$ exceeds $1 - \epsilon$. But then, since ϵ is an arbitrary positive number, we have $|I + T| \geqq 1$. Hence, since $|I| = 1$, (6) follows.

THEOREM 3.2. *With $|\alpha - \beta|$ interpreted as the distance between α and β, $\mathfrak{B}/\mathfrak{T}$ is a complete linear metric space.*

That $\mathfrak{B}/\mathfrak{T}$ is a metric space follows from Theorem 3.1 while its linear properties are evident. To show that it is complete we must prove that, for every sequence $\{\alpha_n\}$ in $\mathfrak{B}/\mathfrak{T}$ such that

$$\lim_{n,m \to \infty} |\alpha_n - \alpha_m| = 0,$$

there exists an element α such that

$$\lim_{n \to \infty} |\alpha_n - \alpha| = 0.$$

Let $\{\alpha_n\}$ be a sequence satisfying the first of these conditions. We choose a subsequence $\{\alpha_{n_k}\}$ such that

$$|\alpha_{n_k} - \alpha_n| \leqq \frac{1}{2^{k+1}}, \qquad \text{for } n \geqq n_k.$$

We then choose an arbitrary element A_{n_1} of α_{n_1} and an element C_1 of $\alpha_{n_2} - \alpha_{n_1}$ such that $|C_1| \leqq \frac{1}{2}$. Setting $A_{n_2} = C_1 + A_{n_1}$, we have A_{n_2} in α_{n_2}

$$|A_{n_2} - A_{n_1}| \leqq \tfrac{1}{2}.$$

[21] For the fact that $\mathfrak{T}$ is divisorless in $\mathfrak{B}$ implies that $\mathfrak{B}/\mathfrak{T}$ contains no two-sided ideals except (0) and $\mathfrak{B}/\mathfrak{T}$ itself. Cf. [19], pp. 56–57 for a discussion of ideals in commutative rings which is readily generalized to cover the case in hand.

Continuing this process we determine a sequence $\{A_{n_k}\}$, with $A_{n_k} \epsilon \alpha_{n_k}$, such that

$$|A_{n_{k+1}} - A_{n_k}| \leqq \frac{1}{2^k}.$$

Thus

$$|A_{n_{k+j}} - A_{n_k}| \leqq \sum_{h=k}^{j-1} |A_{n_{h+1}} - A_{n_h}| \leqq \frac{1}{2^{k-1}}.$$

Hence there exist an element A of $\mathfrak{B}$ such that

$$\lim_{k\to\infty} |A_{n_k} - A| = 0,$$

and if α is the residue class to which A belongs we have, in consequence,

$$\lim_{n\to\infty} |\alpha_n - \alpha| = \lim_{k\to\infty} |\alpha_{n_k} - \alpha| = 0.$$

We note in passing that the space $\mathfrak{B}/\mathfrak{T}$ is non-separable. For every idempotent except 0 in $\mathfrak{B}/\mathfrak{T}$ can be shown to have the norm 1, and if $E(\lambda)$ is the resolution of the identity in $\mathfrak{B}$ of a transformation with spectrum the entire interval $0 \leqq \lambda \leqq 1$, the set $\epsilon(\lambda)$ of idempotents in $\mathfrak{B}/\mathfrak{T}$ such that $\epsilon(\lambda)$ contains $E(\lambda)$ has the property that $\epsilon(\lambda_2) - \epsilon(\lambda_1)$ is an idempotent different from zero if $0 \leqq \lambda_2 < \lambda_1 \leqq 1$.

4. The space $\mathfrak{L}$

We now propose to realize $\mathfrak{B}/\mathfrak{T}$ as an algebraic ring of operators in a certain complex Euclidean space. To define a space $\mathfrak{L}$ suitable for this purpose, we make use of a concept of generalized limit introduced by Banach and Mazur.[22] In the interests of greater generality, however, we shall employ a less restrictive concept of generalized limit than that of these writers, and we begin with a discussion of this concept.[23]

We consider a linear functional defined for all bounded sequences $\{x_n\}$ of real numbers, denoted by $\operatorname{Lim}_{n\to\infty} x_n$, which has the following properties:

(a) $\operatorname{Lim}_{n\to\infty} (x_n + y_n) = \operatorname{Lim}_{n\to\infty} x_n + \operatorname{Lim}_{n\to\infty} y_n$;

(b) $\operatorname{Lim}_{n\to\infty} x_n \geqq 0$, for $x_n \geqq 0$, $n = 1, 2, \cdots$;

(c) $\operatorname{Lim}_{n\to\infty} x_n$ is independent of x_p for each integer p;

(d) $\operatorname{Lim}_{n\to\infty} 1 = 1$.

---

[22] [2], p. 34.

[23] The possibility of generalizing the notion of Banach in this way was pointed out to us by J. v. Neumann; originally we had employed the Banach limit.

For subsequent use, we note that (a) and (b) imply

(e) $$\operatorname*{Lim}_{n\to\infty} x_n \geqq \operatorname*{Lim}_{n} y_n, \quad \text{if} \quad x_n \geqq y_n \quad \text{for} \quad n = 1, 2, \cdots.$$

The reader will observe that the four preceding conditions differ from the four basic properties of the Banach limit, as given in the reference cited, in the following respects: first, we do not require homogeneity; second, and more important, the Banach limit has the property

(1) $$\operatorname*{Lim}_{n\to\infty} x_{n+1} = \operatorname*{Lim}_{n\to\infty} x_n$$

in place of our (c). Since from (1) one has

$$\operatorname*{Lim}_{n\to\infty} x_{n+p+1} = \operatorname*{Lim}_{n\to\infty} x_n,$$

it is clear that (1) implies (c).[24]

We now wish to show that homogeneity is a consequence of conditions (a) – (d) and that the use of (c) instead of (1) does not affect the other essential properties of $\operatorname*{Lim}_{n\to\infty} x_n$.

To begin we observe that (a) implies

(2) $$\operatorname*{Lim}_{n\to\infty} rx_n = r \operatorname*{Lim}_{n\to\infty} x_n \text{ for all rational } r.$$

We now consider an arbitrary bounded sequence $\{x_n\}$, and rational upper and lower bounds, R and r, respectively, of $\{x_n\}$. Invoking (2), (d), and (e), we then obtain

$$R = \operatorname*{Lim}_{n\to\infty} R \geqq \operatorname*{Lim}_{n\to\infty} x_n \geqq \operatorname*{Lim}_{n\to\infty} r = r.$$

Hence, since R is any rational upper bound of $\{x_n\}$, and r any rational lower bound, we have

(3) $$\text{l. u. b. } x_n \geqq \operatorname*{Lim}_{n\to\infty} x_n \geqq \text{g. l. b. } x_n$$

Moreover, if we now invoke (c) in conjunction with (3), it becomes clear at once that we must have

(f) $$\limsup_{n\to\infty} x_n \geqq \operatorname*{Lim}_{n\to\infty} x_n \geqq \liminf_{n\to\infty} x_n$$

Thus, we have the important property

(g) $$\operatorname*{Lim}_{n\to\infty} x_n = \lim_{n\to\infty} x_n \quad \text{whenever} \quad \lim_{n\to\infty} x_n \text{ exists.}$$

[24] The argument of Banach, loc. cit., thus serves to establish the existence of a functional with the properties (a)–(d). An elegant and simple direct proof of the existence of such functionals has been obtained by Ulam and Kakutani independently, but has not been published. Moreover, it is not difficult to show that there exist functionals satisfying (a)–(d) but not (1). This latter fact, however, is not of essential importance in the present paper, but rather in certain related investigations. Added in proof: Since the completion of this paper J. v. Neumann has developed a general theory of limits of the sort used here. His results will appear in a forthcoming number of the ANNALS OF MATHEMATICS STUDIES dealing with the theory of measure.

Furthermore, since it is now clear that $\operatorname{Lim}_{n\to\infty}$ is a bounded additive functional on the space of all bounded sequences, we can conclude that it is homogeneous:

(h) $$\operatorname*{Lim}_{n\to\infty} ax_n = a \operatorname*{Lim}_{n\to\infty} x_n, \quad \text{for all numbers } a.$$

We now extend the notion of generalized limit to bounded sequences of complex numbers in the obvious way; if $\{x_n\}$ is such a sequence we set

(i) $$\operatorname*{Lim}_{n\to\infty} x_n = \operatorname*{Lim}_{n\to\infty} \Re x_n + i \operatorname*{Lim}_{n\to\infty} \Im x_n.$$

It is then readily proved that properties (a), (c), (d), (g), (h), persist in the complex case.

Hereafter, as occasion requires, we shall refer to the properties of $\operatorname{Lim}_{n\to\infty} x_n$ as given above by letter without other comment.

We turn now to the construction of our space $\mathfrak{L}$. First, we consider the class $\mathfrak{L}''$ of all sequences $\{f_n\}$ in Hilbert space $\mathfrak{H}$ which are weakly convergent to zero; this class is evidently a module when we define addition and scalar multiplication by the equations

$$\{f_n\} + \{g_n\} = \{f_n + g_n\},$$
$$a\{f_n\} = \{af_n\}.$$

In $\mathfrak{L}''$, we define $(\{f_n\}, \{g_n\})$ by the equation

$$(\{f_n\}, \{g_n\}) = \operatorname*{Lim}_{n\to\infty} (f_n, g_n),$$

invoking the boundedness of the sequences $\{f_n\}$ and $\{g_n\}$ to assure the boundedness of the sequence of numbers $\{(f_n, g_n)\}$. Then from the properties of $\operatorname{Lim}_{n\to\infty}$ given above and the properties of the inner product in $\mathfrak{H}$, we have

$$(a\{f_n\} + b\{g_n\}, \{h_n\}) = a(\{f_n\}, \{h_n\}) b(\{g_n\}, \{h_n\}),$$
$$(\{f_n\}, \{g_n\}) = \overline{(\{g_n\}, \{f_n\})},$$
$$(\{f_n\}, \{f_n\}) \geqq 0.$$

Thus $(\{f_n\}, \{g_n\})$ has all the properties requisite for an inner product in $\mathfrak{L}''$ except that requiring that $(\{f_n\}, \{f_n\}) = 0$ if and only if $\{f_n\} = 0$, and it is easy to see that this requirement is not fulfilled, since $\mathfrak{L}''$ contains sequences strongly convergent to zero. However, as A. E. Taylor [18] has pointed out, this requirement is not an essential one, since we can regard it not as a postulate but as a definition of zero and thus of equality. In our case, this means that we must identify the sequences $\{f_n\}$ and $\{g_n\}$ provided

$$\operatorname*{Lim}_{n\to\infty} |f_n - g_n|^2 = 0.$$

In this way, we obtain a class $\mathfrak{L}'$ of elements $f, g, \cdots$, the quotient group of the additive group $\mathfrak{L}''$ by the subgroup of elements $\{f_n\}$ such that

$$\operatorname*{Lim}_{n\to\infty} |f_n|^2 = 0.$$

If under this homomorphism $\{f_n\} \to f$, $\{g_n\} \to g$, we define

$$(f, g) = (\{f_n\}, \{g_n\}).$$

Thus we achieve in $\mathfrak{L}'$ a (possibly incomplete) complex Euclidean space; that is to say, $\mathfrak{L}'$ is a module and the function (f, g) has the properties

$$(af + bg, h) = a(f, h) + b(g, h),$$

$$(f, g) = \overline{(g, f)},$$

$$(f, f) \geqq 0,$$

$$(f, f) = 0 \quad \text{implies} \quad f = 0.$$

Since the pertinent facts in this connection are discussed in the paper of Taylor cited above, it is unnecessary for us to dwell on them here.

It now remains for us to consider the space $\mathfrak{L}'$ with reference to the matters of completeness and separability. The answers to both questions are provided through the two simple lemmas which follow.

LEMMA 4.1. *The cardinal number of $\mathfrak{L}'$ is not greater than the cardinal number c of the continuum.*

Let $\{\varphi_m\}$ be an arbitrary complete orthonormal set in $\mathfrak{H}$, $\{f_n\}$ an arbitrary element of $\mathfrak{L}''$. Then, if

$$f_n = \sum_{m=1}^{\infty} a_{n,m}\varphi_m$$

$\{a_{n,m}\}$ is a bounded matrix. Thus the cardinal number of $\mathfrak{L}''$ does not exceed the cardinal number of the class of all bounded infinite matrices, and the cardinal number of the latter class is c. Hence the cardinal number of $\mathfrak{L}'$ is certainly less than c.

LEMMA 4.2. *The space $\mathfrak{L}'$ contains an orthonormal set with cardinal number* c.

We consider an enumeration $\{r_n\}$ of the positive rational numbers, a complete orthonormal set $\{\varphi_n\}$ in $\mathfrak{H}$, and the correspondence $r_n \leftrightarrow \varphi_n$ between them. We denote by $\{\{\varphi_n^\alpha\}\}$ the class of all infinite subsequences of $\{\varphi_n\}$, α running over some set which we leave undesignated. Now let a_1 and a_2 be any two distinct positive numbers, $\{r_n'\}$ and $\{r_n''\}$ infinite subsequences of $\{r_n\}$ convergent to a_1 and a_2 respectively. Corresponding to $\{r_n'\}$ and $\{r_n''\}$ we have two subsequences of $\{\varphi_n\}$ which belong to $\{\{\varphi_n^\alpha\}\}$; we denote these by $\{\varphi_n^{\alpha_1}\}$ and $\{\varphi_n^{\alpha_2}\}$, respectively. Then, for n larger than some integer N, we have $\varphi_n^{\alpha_1} \neq \varphi_n^{\alpha_2}$ and hence

$$\operatorname*{Lim}_{n\to\infty} (\varphi_n^{\alpha_1}, \varphi_n^{\alpha_2}) = 0.$$

Moreover,

$$\operatorname*{Lim}_{n\to\infty} |\varphi_n^\alpha|^2 = 1,$$

for all α. Hence, if φ_α denotes the element of $\mathfrak{L}'$ containing $\{\varphi_n^\alpha\}$, it follows that the set of all $\{\varphi_\alpha\}$ contains an orthonormal set with cardinal number $\mathfrak{c}$.

THEOREM 4.1. *The space $\mathfrak{L}'$ is incomplete.*

Consider an orthonormal set $\{\psi_\alpha\}$ in $\mathfrak{L}'$ with cardinal number at least $\mathfrak{c}$. Then the set of all $\sum_\alpha a_\alpha\psi_\alpha$ with $\sum_\alpha |a_\alpha|^2 < \alpha$ has cardinal number at least $2^{\mathfrak{c}}$, and hence, in view of Lemma 4.1, this set cannot belong to $\mathfrak{L}'$. Therefore $\mathfrak{L}'$ is incomplete.[25]

We now denote by $\mathfrak{L}$ the space obtained by completing $\mathfrak{L}'$; the details of this construction are described in [5] and in [13], so we need not consider them here.

THEOREM 4.2. *The dimension number of $\mathfrak{L}$ is* $\mathfrak{c}$.

Since $\mathfrak{L}'$ is dense in $\mathfrak{L}$, the dimension number of $\mathfrak{L}$ cannot exceed $\mathfrak{c}$, by Lemma 4.1. But by Lemma 4.2, it cannot be less than $\mathfrak{c}$.

5. The algebraic ring $\mathfrak{M}$ and congruence modulo $\mathfrak{T}$ in $\mathfrak{B}$

We now consider transformations induced in the space $\mathfrak{L}$ by means of members of the ring $\mathfrak{B}$.

LEMMA 5.1. *Let A be an arbitrary bounded everywhere-defined transformation in $\mathfrak{H}$, $\{f_n\}$ an arbitrary sequence of the class $\mathfrak{L}''$. Then $\{Af_n\}$ is in $\mathfrak{L}''$ and* $\operatorname*{Lim}_{n\to\infty} |Af_n|^2 = 0$ *if* $\operatorname*{Lim}_{n\to\infty} |f_n|^2 = 0$.

That $\{Af_n\}$ is in $\mathfrak{L}''$ follows at once from the fact that a bounded transformation is weakly continuous. And since we have $|Af_n|^2 \leqq |A|^2|f_n|^2$, $n = 1, 2, \cdots$, it follows from property (e) of Lim that $\operatorname*{Lim}_{n\to\infty} |f_n|^2 = 0$ implies

$$\operatorname*{Lim}_{n\to\infty} |Af_n|^2 = 0.$$

THEOREM 5.1. *Let A be an arbitrary member of $\mathfrak{B}$. Then, if f is an arbitrary element of $\mathfrak{L}'$, and $\{f_n\}$ belongs to f, we set*

$$g = T_1(A)f,$$

where g is the element of $\mathfrak{L}'$ containing $\{Af_n\}$. The transformation $T_1(A)$ so defined in $\mathfrak{L}'$ is a single valued linear bounded transformation with bound not exceeding the bound of A in $\mathfrak{H}$. Thus $T_1(A)$ has a unique closed bounded extension $T(A)$ with domain $\mathfrak{L}$, and the bound of $T(A)$ does not exceed the bound of A in $\mathfrak{H}$.

That $T_1(A)$ is single-valued follows at once from Lemma 5.1, while its linear character is a consequence of the linearity of A. Since, in addition,

$$|T_1(A)f|^2 = \operatorname*{Lim}_{n\to\infty} |Af_n|^2 \leqq \operatorname*{Lim}_{n\to\infty} |A|^2|f_n|^2 = |A|^2|f|^2$$

[25] This simple proof of Theorem 4.1 was suggested by J. v. Neumann. The theorem can also be proved directly.

it is evident that $T_1(A)$ is bounded with bound less than or equal to $|A|$. Thus the transformation $\bar{T}_1(A) = T(A)$ exists and has domain $\mathfrak{L}$, while its bound is clearly the same as that of $T_1(A)$.

THEOREM 5.2. *The class $\mathfrak{M}$ of operators $T(A)$ in $\mathfrak{L}$, defined for all A in $\mathfrak{B}$, is an algebraic ring of operators in the class of all bounded everywhere defined operators in $\mathfrak{L}$, and is a homomorphism of $\mathfrak{B}$ with respect to the operations $+,\cdot,^*$, and scalar multiplication; that is,*

$$T(A + B) = T(A) + T(B), \qquad T(AB) = T(A)T(B),$$
$$T(A^*) = T^*(A), \qquad T(aA) = aT(A).$$

All of these relations are quite obvious, except possibly $T(A^*) = T^*(A)$. To prove this we consider two arbitrary elements f and g of $\mathfrak{L}'$ and sequences $\{f_n\}$ and $\{g_n\}$ belonging to f and g, respectively. Then

$$(T(A)f, g) = \operatorname*{Lim}_{n\to\infty} (Af_n, g_n) = \operatorname*{Lim}_{n\to\infty} (f_n, A^*g_n) = (f, T(A^*)g).$$

Thus $T(A^*)$ and $T^*(A)$ coincide on $\mathfrak{L}'$, and therefore throughout $\mathfrak{L}$.

LEMMA 5.2. *A necessary and sufficient condition that $T(A)$ be the transformation in $\mathfrak{L}$ which takes every element of $\mathfrak{L}$ into zero is that A belong to the ideal $\mathfrak{T}$ of totally continuous operators in $\mathfrak{B}$. Thus $T(A) = T(B)$ if and only if A is congruent to B modulo $\mathfrak{T}$.*

From the homomorphism $\mathfrak{B} \to \mathfrak{M}$, it follows that the set of all A in $\mathfrak{B}$ such that $T(A) = 0$ is a two-sided ideal $\mathfrak{J}$ in $\mathfrak{B}$. Moreover, since a totally continuous transformation A in $\mathfrak{B}$ takes weakly convergent sequences into strongly convergent ones, it follows from property (e) of $\operatorname*{Lim}_{n\to\infty}$ that $\mathfrak{J} \supseteq \mathfrak{T}$. Hence, by Theorem 1.4, we have either $\mathfrak{J} = \mathfrak{T}$ or $\mathfrak{J} = \mathfrak{B}$. But since $\mathfrak{J}$ clearly fails to contain the identity in $\mathfrak{B}$, we must conclude that $\mathfrak{J} = \mathfrak{T}$, which establishes the lemma.

Since it also follows immediately from the homomorphism $\mathfrak{B} \to \mathfrak{M}$ that $\mathfrak{M}$ is isomorphic to the ring $\mathfrak{B}/\mathfrak{J}$, where $\mathfrak{J}$ is the ideal of all A in $\mathfrak{B}$ such that $T(A) = 0$, we can now conclude that $\mathfrak{M}$ is isomorphic to $\mathfrak{B}/\mathfrak{T}$. More precisely, we have

THEOREM 5.3. *The algebraic ring $\mathfrak{M}$ is isomorphic to $\mathfrak{B}/\mathfrak{T}$ with respect to the operations $+$, $\cdot$, *, $a\cdot$, an element $T(A)$ of $\mathfrak{M}$ corresponding to the element α of $\mathfrak{B}/\mathfrak{T}$ if and only if A belongs to α.*

DEFINITION 5.1. *If α is an arbitrary element of $\mathfrak{B}/\mathfrak{T}$, we define $T(\alpha)$ as the element of $\mathfrak{M}$ corresponding to α under the isomorphism of Theorem 5.3.*

Evidently $T(\alpha)$ is identical with $T(A)$, for all A in α, and we shall continue to use both notations for elements of $\mathfrak{M}$ as occasion requires.

THEOREM 5.4. *An element $T(\alpha)$ of $\mathfrak{M}$ is self-adjoint, partially isometric, or unitary, respectively if and only if α has that property in the sense of Definition 2.1. An element $T(\alpha)$ of $\mathfrak{M}$ is a projection if and only if α is an idempotent according to that definition.*

The assertion of the theorem concerning self-adjointness is obvious. To prove the other parts of the theorem we note first that since the properties which form

the various criteria of Definition 2.1 are all defined in terms of the operations · and *, it follows from Theorem 5.3 that an element α of $\mathfrak{B}/\mathfrak{T}$ possesses one of them if and only if $T(\alpha)$ does. But for transformations, each of these properties is characteristic of the class of transformations in question: more precisely, an everywhere defined bounded linear operator T in $\mathfrak{L}$ is a projection if and only if $T^2 = T^* = T$,[26] is partially isometric if and only if T^*T is a projection,[27] and thus is obviously unitary if and only if T^*T and TT^* are equal to the identity in $\mathfrak{L}$. Hence the theorem follows.

The reader will note that Theorems 2.3–2.7 can be interpreted now to yield assertions concerning the homomorphism $\mathfrak{B} \to \mathfrak{M}$; the details here are obvious and we omit them.

THEOREM 5.5. *Let α be an arbitrary element of $\mathfrak{B}/\mathfrak{T}$. Then the bound of the operator $T(\alpha)$ in $\mathfrak{L}$ is $|\alpha|$; in other words, the isomorphism $\mathfrak{B}/\mathfrak{T} \leftrightarrow \mathfrak{M}$ is an isometry. Thus $\mathfrak{M}$ is closed in the uniform topology for operators.*

Evidently the concluding assertion is a consequence of the first one and Theorem 3.2. Hence we need only show $|T(\alpha)| = |\alpha|$.

From the final statement of Theorem 5.1 and the definition of the norm in $\mathfrak{B}/\mathfrak{T}$, we have at once

$$|T(\alpha)| \leqq |\alpha|. \tag{1}$$

Hence we need only establish

$$|T(\alpha)| \geqq |\alpha|. \tag{2}$$

To prove (2), we first select an arbitrary element A of α, with canonical decomposition $A = WB$. We then denote by λ the lowest upper bound of those points of the spectrum of B which are either limit points of the spectrum or characteristic values of infinite multiplicity, and by S the set of points μ in the spectrum of B such that μ exceeds λ. Then S clearly consists entirely of isolated points, each a characteristic value of finite multiplicity. Furthermore, either S is a finite sequence $\{\mu_n\}$ or an infinite sequence with λ as limit. Hence if $\mathfrak{M}_n$ is the characteristic manifold of B corresponding to μ_n, $n = 1, 2, \cdots$, and we set $C = B - \lambda I$ on $\mathfrak{M} = \sum_n \mathfrak{M}_n$, $C = 0$ on $\mathfrak{H} \ominus \mathfrak{M}$, C belongs to $\mathfrak{T}$. Thus, if $B_1 = B - C$, then $A_1 = WB_1$ belongs to α. Moreover, this is evidently the canonical decomposition of A_1, so $|A_1| = |B_1|$. Hence we have

$$|B_1| \geqq |\alpha|. \tag{3}$$

Now let us consider the transformation $T(B_1)$ in $\mathfrak{L}$. We distinguish two cases, according as the sequence $\{\mu_n\}$ is infinite or finite. If $\{\mu_n\}$ is infinite, $\mathfrak{M} = \sum_n \mathfrak{M}_n$ is a Hilbert space and contains an infinite orthonormal set $\{\varphi_n\}$. Furthermore, $B_1\varphi_n = \lambda\varphi_n$, $n = 1, 2, \cdots$. Thus, if φ is the element of $\mathfrak{L}'$ to which

[26] [17], Theorems 2.35, 2.36.

[27] [9], Lemma 4.3.2.

$\{\varphi_n\}$ belongs, we have

$$(T(B_1)\varphi, \varphi) = \lim_{n\to\infty} (B_1\varphi_n, \varphi_n) = \lambda.$$

$$(\varphi, \varphi) = \lim_{n\to\infty} (\varphi_n, \varphi_n) = 1. \tag{4}$$

But clearly $\lambda = |B_1|$, and hence we have

$$|T(B_1)| \geqq |B_1|. \tag{5}$$

Now suppose $\{\mu_n\}$ is finite, and let $E(\lambda)$ be the resolution of the identity for B_1 in $\mathfrak{H}$. Then, since λ is a limit point of the spectrum of B_1, there exists a monotone increasing sequence $\{\lambda_n\}$ with limit λ, such that $E(\lambda_{n+1}) - E(\lambda_n)$ is different from zero, $n = 1, 2, \cdots$. Hence we can select an orthonormal set $\{\varphi_n\}$ with φ_n in the range of $E(\lambda_{n+1}) - E(\lambda_n)$, $n = 1, 2, \cdots$. Moreover, for every n, we have $\lambda_n \leqq (B_1\varphi_n, \varphi_n) \leqq \lambda_{n+1}$. Hence if φ is the element of $\mathfrak{F}'$ containing $\{\varphi_n\}$, we have (4) in this case also.

Finally, since $B_1 = W^*A_1$ we have $T(B_1) = T(W^*)T(A_1)$ and since $T(W^*)$ is partially isometric by Theorems 5.4 and 2.6, it follows that we have $|T(W^*)| = 1$, and hence that

$$|T(A_1)| \geqq |T(B_1)| \tag{6}$$

holds. But $T(A_1) = T(\alpha)$, and hence from (3), (5) and (6) we have (2) which completes the proof of the theorem.

We now wish to prove that $\mathfrak{M}$ is not closed in the weak or strong topologies for operators. The proof reposes on the following lemma concerning monotone sequences of projections in $\mathfrak{M}$.

LEMMA 5.3. *Let $\{T(\epsilon_n)\}$ be a sequence of projections in $\mathfrak{M}$ such that $T(\epsilon_{n+1}) \leqq T(\epsilon_n)$, $n = 1, 2, \cdots$. Then, if $\lim_{n\to\infty} T(\epsilon_n) = 0$, $T(\epsilon_n) = 0$ for all n greater than some integer M.*

By Theorem 5.4, each of the terms of $\{\epsilon_n\}$ is an idempotent, and consequently by Theorem 2.4, contains a projection E_n of $\mathfrak{B}$. The sequence $\{E_n\}$, however, is evidently not necessarily monotone, and our next step is to show that there exists a monotone non-increasing sequence $\{F_n\}$ of projections in $\mathfrak{B}$ such that $E_n - F_n$ is in $\mathfrak{T}$, $n = 1, 2, \cdots$. We begin by setting $F_1 = E_1$ and then, assuming that F_n is determined for all $n \leqq N$, we show that F_{N+1} can be defined.

We note first that

$$T(\epsilon_n)T(\epsilon_{n+1})T(\epsilon_n) = T(\epsilon_{n+1})$$

and hence that $F_NE_{N+1}F_N - E_{N+1}$ is in $\mathfrak{T}$. Consequently, it follows that $(F_NE_{N+1}F_N)^2 - F_NE_{N+1}F_N$ is in $\mathfrak{T}$. Thus, if $\mathfrak{M}_N$ is the range of F_N, $F_NE_{N+1}F_N$ induces in $\mathfrak{M}_N$ a self-adjoint transformation E'_{N+1} congruent to its square modulo the class of all totally continuous operators in $\mathfrak{M}_N$. Hence, by Theorem 2.4, there exists in $\mathfrak{M}_N$ a projection F'_{N+1} congruent to E'_{N+1} modulo that class. Therefore if F_{N+1} is the projection in $\mathfrak{H}$ which is equal to F'_{N+1} in $\mathfrak{M}_N$, equal to

zero in $\mathfrak{H} \ominus \mathfrak{M}_N$, then $F_{N+1} - F_N E_{N+1} F_N$ is in $\mathfrak{T}$. But then $F_{N+1} - E_{N+1}$ is in $\mathfrak{J}$ and since F_{N+1} is equal to zero in $\mathfrak{H} \ominus \mathfrak{M}_N$, we have $F_{N+1} \leqq F_N$. Thus the sequence $\{F_n\}$ with the stated properties exists.

Now let us suppose that $T(\epsilon_n)$ is never zero. Then, if $\lim_{n\to\infty} T(\epsilon_n) = 0$, $T(\epsilon_n)$ must be different from $T(\epsilon_{n+1})$ for an infinite number of values of n. Hence we can select a subsequence $\{T(\epsilon_{n_k})\}$ such that $T(\epsilon_{n_{k+1}}) < T(\epsilon_{n_k})$ holds, $k = 1, 2, \cdots$. Consequently, we clearly have $F_{n_{k+1}} < F_{n_k}$, $k = 1, 2, \cdots$, since $T(\epsilon_n) = T(F_n)$. Therefore, if $\mathfrak{N}_k$ is the range of F_{n_k}, none of the spaces $\mathfrak{N}_k \ominus \mathfrak{N}_{k+1}$ is empty, and we can select an orthonormal set $\{\varphi_k\}$ with φ_k in $\mathfrak{N}_k \ominus \mathfrak{N}_{k+1}$, $k = 1, 2, \cdots$. Let $\mathfrak{N}_0$ be the closed linear manifold determined by $\{\varphi_k\}$, F_0 the projection with range $\mathfrak{N}_0$. Then $F_0 - F_0 F_{n_k} F_0$ is the projection with range determined by $\{\varphi_j\}$, $j = 1, 2, \cdots k$, and hence belongs to $\mathfrak{T}$. Consequently $T(F_0)T(F_{n_k})T(F_0) = T(F_0)$ and $T(F_0)$ is clearly not zero. Hence, since we obviously have

$$\lim_{k\to\infty} T(F_0)T(F_{n_k})T(F_0) = 0$$

if $\lim_{n\to\infty} T(\epsilon_n) = 0$, it follows that the latter is impossible. Therefore $T(\epsilon_n) = 0$ for all n sufficiently large, as we wished to show.

It is of some interest to note the following alternative statement of the preceding lemma; if $\{T(\epsilon_n)\}$ is an infinite sequence of orthogonal projections in $\mathfrak{M}$, and if $\sum_{n=1}^{\infty} T(\epsilon_n)$ is the identity in $\mathfrak{L}$, then all but a finite number of terms of the sequence $T(\epsilon_n)$ are zero.

THEOREM 5.6. *The algebraic ring $\mathfrak{M}$ is not closed in either the weak or strong topology for operators.*

We consider a monotone sequence $\{T(\epsilon_n)\}$ of projections in $\mathfrak{M}$, with $T(\epsilon_{n+1}) > T(\epsilon_n)$, $n = 1, 2, \cdots$. Such a sequence is readily generated by means of a sequence $\{E_n\}$ of projections in $\mathfrak{B}$ such that $E_{n+1} - E_n$ is a projection with range a Hilbert space, $n = 1, 2, \cdots$; we have only to choose ϵ_n as the residue class of E_n. The sequence $T(\epsilon_n)$, being monotone, is convergent in both the weak and the strong topologies and has a limit which is a projection. Let us suppose that this limit is in $\mathfrak{M}$ and hence that it has the form $T(\epsilon)$ where ϵ is an idempotent in $\mathfrak{B}/\mathfrak{T}$. Then $\{T(\epsilon_n) - T(\epsilon)\}$ is a monotone non-increasing sequence of projections in $\mathfrak{M}$ with limit zero and thus $T(\epsilon_n) = T(\epsilon)$ for all n greater than some integer N, by Lemma 5.3. This, however, is impossible in view of the fact that $\{T(\epsilon_n)\}$ was chosen with $T(\epsilon_{n+1}) > T(\epsilon_n)$, $n = 1, 2, \cdots$; we must conclude therefore that the sequence $\{T(\epsilon_n)\}$ has no limit in $\mathfrak{M}$.

We proceed now to examine the relationship between the spectrum of a self-adjoint member of $\mathfrak{M}$ and the corresponding self-adjoint transformations in $\mathfrak{B}$. We begin with a formal definition.

DEFINITION 5.2. *Let A be an arbitrary self-adjoint transformation in $\mathfrak{H}$. Then a point λ of the spectrum of A which is a limit point of the spectrum or a character-*

istic value of infinite multiplicity is called a point of condensation of the spectrum.[28] *The set of all such points is called the condensed spectrum of A. The complementary set in the λ-plane is called the augmented resolvent set of A.*

THEOREM 5.7. *Let A be a self-adjoint transformation in $\mathfrak{H}$. Then a necessary and sufficient condition that λ belong to the augmented resolvent set of A is that the manifold $\mathfrak{M}$ of solutions of the equation $Af - \lambda f = 0$ have a finite dimension number and that in $\mathfrak{H} \ominus \mathfrak{M}$, $A - \lambda I$ induce a transformation with bounded inverse.*

Let $\mathfrak{M}$ be the manifold of zeros of A, A_1 the transformation induced in $\mathfrak{H} \ominus \mathfrak{M}$ by A. Then the condition of the theorem may be stated in this way: $\mathfrak{M}$ is finite-dimensional and λ is in the resolvent set of A_1. But the latter is possible if and only if λ is a finite distance from the spectrum of A_1, which is to say that λ is an isolated point of the spectrum of A or belongs to the resolvent set of A. Thus the theorem follows.

THEOREM 5.8. *Let A be a self-adjoint transformation in $\mathfrak{H}$. Then the resolvent set of the transformation $T(A)$ in $\mathfrak{L}$ is the augmented resolvent set of A and the spectrum of $T(A)$ is the condensed spectrum of A. Every point in the spectrum of $T(A)$ is a characteristic value with multiplicity* c.

Let λ belong to the augmented resolvent set of A, and let $\mathfrak{M}$ be the manifold of zeros of $A - \lambda I$, E the projection operator of $\mathfrak{H} \ominus \mathfrak{M}$. Let B be equal in $\mathfrak{M}$ to zero, and in $\mathfrak{H} \ominus \mathfrak{M}$ to the inverse of the transformation induced in that space by $A - \lambda I$. Then B is in $\mathfrak{B}$ and $B(A - \lambda I) = E$. Thus $T(B)T(A - \lambda I) = T(E)$. But $T(E)$ is the identity in $\mathfrak{L}$, since $I - E$ is in $\mathfrak{T}$, and hence $T(B) = [T(A) - \lambda T(I)]^{-1}$. Therefore λ is in the resolvent set of $T(A)$.

Now suppose λ is in the condensed spectrum of A. Then λ is either a characteristic value of infinite multiplicity or a limit point of the spectrum of A, and in either case we can select an orthonormal set $\{\varphi_n\}$ in $\mathfrak{H}$ such that

$$\lim_{n\to\infty} (A\varphi_n, \varphi_n) = \lambda$$

$$\lim_{n\to\infty} (A\varphi_n, A\varphi_n) = \lambda^2.$$

Thus

$$\lim_{n\to\infty} |A\varphi_n - \lambda\varphi_n|^2 = \lim_{n\to\infty} (|A\varphi_n|^2 + \lambda^2|\varphi_n|^2 - 2\lambda(A\varphi_n, \varphi_n)) = 0.$$

But then, if $\{\psi_n\}$ is any subsequence of $\{\varphi_n\}$ and ψ the element of $\mathfrak{L}'$ containing $\{\psi_n\}$, we have

$$T(A)\psi = \lambda\psi.$$

Moreover, there exist c such subsequences such that any two have only a finite number of terms in common and consequently c orthogonal elements of $\mathfrak{L}'$ satisfying the preceding equation. Therefore λ is a characteristic value of $T(A)$ with multiplicity c.

[28] These are the Häufungspunkte of the spectrum in the sense of Weyl, [20].

We have now shown that every point of the augmented resolvent set of A is in the resolvent set of $T(A)$ and that every point of the condensed spectrum of A is in the spectrum of $T(A)$. But since the augmented resolvent set and the condensed spectrum of A together constitute the entire λ-plane, they must be respectively the entire resolvent set and the entire spectrum of A.

It should be observed that we cannot infer from theorem 5.8 that the sum of the characteristic manifolds of $T(A)$ is $\mathfrak{L}$. Whether or not this is so we are unable to say at present.

We now have from Theorem 5.8 the classical theorem of Weyl [20].

THEOREM 5.9. *Let A and B be two self-adjoint transformations in $\mathfrak{H}$, such that $A - B$ is totally continuous. Then A and B have the same condensed spectrum and the same augmented resolvent set.*[29]

The theorem is obvious since $T(A) = T(B)$.

We proceed now to determine the resolution of the identity corresponding to a self-adjoint transformation in $\mathfrak{M}$ by means of the resolution of the identity of a corresponding member of $\mathfrak{B}$. If A is self-adjoint in $\mathfrak{H}$, $E(\lambda)$ its resolution of the identity, it is clear that $T(E(\lambda))$ has many of the properties of a resolution of the identity in $\mathfrak{L}$. Specifically, it is readily shown that the following assertions hold:

(1) $T(E(\lambda))$ permutes with $T(A)$;

(2) $T(E(\lambda))T(E(\mu)) = T(E(\mu))\ T(E(\lambda)) = T(E(\mu))$ for $\mu \leqq \lambda$;

(3) $T(E(\lambda)) = 0$ if λ is less than the lower bound of $T(A)$, $T(E(\lambda)) = 1$ if λ exceeds the upper bound of $T(A)$;

(4) in the range of $T(E(\lambda))$, the upper bound of $T(A)$ does not exceed λ, and in the range of $1 - T(E(\lambda))$ the lower bound of $T(A)$ is not less than λ.

In spite of these facts $T(E(\lambda))$ fails usually to be the resolution of the identity for $T(A)$. This is readily seen in view of Lemma 5.3, which assures us that in general we do not have

$$\lim_{\epsilon \to 0} T(E(\lambda + \epsilon)) = T(E(\lambda)).$$

Moreover, if A and B are self-adjoint and congruent modulo $\mathfrak{T}$, $E(\lambda)$ and $F(\lambda)$ their respective resolutions of the identity, we do not in general have $E(\lambda) - F(\lambda)$ in $\mathfrak{T}$. Thus for a self-adjoint transformation $T(A)$ in $\mathfrak{M}$ we can exhibit many monotone families of projections with the properties (1) – (4) above, none of which is the resolution of the identity of A.

We can however, derive from any one of these families of projections the resolution of the identity belonging to the member of $\mathfrak{M}$ in question. The procedure is described in the following theorem.

THEOREM 5.10. *Let A be a self-adjoint transformation in $\mathfrak{H}$, $E(\lambda)$ its resolution of the identity. Let $\boldsymbol{E}(\lambda)$ be the family of projections in $\mathfrak{L}$ defined by the equation*

$$\boldsymbol{E}(\lambda) = \lim_{\epsilon \to 0} T(E(\lambda + \epsilon)), \qquad \epsilon > 0.$$

[29] Weyl proves also that if A is an arbitrary bounded self-adjoint transformation, there exists a totally continuous self-adjoint transformation T such that $A + T$ has a pure point spectrum. At present we see no direct way to derive this result from ours. Cf. also [14].

Then $\boldsymbol{E}(\lambda)$ is the resolution of the identity in $\mathfrak{L}$ of $T(A)$.

The existence of the limit $\boldsymbol{E}(\lambda)$ follows of course from the monotone character of the family $T(\boldsymbol{E}(\lambda))$ which follows in turn from the corresponding property of $E(\lambda)$. We shall now show that $\boldsymbol{E}(\lambda)$ has the following six properties:

(1) $\boldsymbol{E}(\lambda)$ permutes with $T(A)$;

(2) $\boldsymbol{E}(\lambda)\boldsymbol{E}(\mu) = \boldsymbol{E}(\mu)\boldsymbol{E}(\lambda) = \boldsymbol{E}(\mu)$ for $\mu \leqq \lambda$;

(3) $\boldsymbol{E}(\lambda) = 0$ for $\lambda < a$ (a the lower bound of $T(A)$), and $\boldsymbol{E}(\lambda) = 1$ for $\lambda \geqq b$ (b the upper bound of $T(A)$);

(4) $\lim_{\epsilon \to 0} \boldsymbol{E}(\lambda + \epsilon) = \boldsymbol{E}(\lambda)$, $\epsilon > 0$;

(5) in the range of $\boldsymbol{E}(\lambda)$, the upper bound of $T(A)$ does not exceed λ;

(6) in the range of $1 - \boldsymbol{E}(\lambda)$, the lower bound of $T(A)$ is not less than λ and if it is equal to λ it is not attained.

The validity of (1) follows at once from the permutability of $T(A)$ and $T(\boldsymbol{E}(\lambda))$.

To prove (2), we note first that

$$T(\boldsymbol{E}(\lambda + \epsilon))T(\boldsymbol{E}(\mu)) = T(\boldsymbol{E}(\mu))T(\boldsymbol{E}(\lambda + \epsilon)) \quad \text{for all} \quad \epsilon > 0.$$

Thus, allowing ϵ to tend to zero, we have

$$\boldsymbol{E}(\lambda)T(\boldsymbol{E}(\mu)) = T(\boldsymbol{E}(\mu))\boldsymbol{E}(\lambda).$$

Hence

$$\boldsymbol{E}(\lambda)T(\boldsymbol{E}(\mu + \epsilon)) = T(\boldsymbol{E}(\mu + \epsilon))\boldsymbol{E}(\lambda)$$

for all μ and all $\epsilon > 0$. Consequently, again allowing ϵ to approach zero, we have

$$\boldsymbol{E}(\lambda)\boldsymbol{E}(\mu) = \boldsymbol{E}(\mu)\boldsymbol{E}(\lambda).$$

Moreover, if $\mu \leqq \lambda$, we have clearly $\boldsymbol{E}(\mu) \leqq \boldsymbol{E}(\lambda)$, and so $\boldsymbol{E}(\lambda)\boldsymbol{E}(\mu) = \boldsymbol{E}(\mu)$. Therefore (2) holds.

Now let a be the lower bound of $T(A)$. Then by Theorem 5.8, a is the lower bound of the condensed spectrum of A and it follows therefore that if λ is less than a, then the range of $\boldsymbol{E}(\lambda)$ must be finite-dimensional. For otherwise the spectrum of A would have points of condensation less than or equal to λ. Thus $T(\boldsymbol{E}(\lambda)) = 0$ for $\lambda < a$ and therefore $\boldsymbol{E}(\lambda) = 0$ for $\lambda < a$.

On the other hand, let b be the upper bound of the spectrum of $T(A)$. Then b is the upper bound of the condensed spectrum of A and by a similar argument $I - \boldsymbol{E}(\lambda)$ has a finite-dimensional range, $\lambda \geqq b$. Thus $T(\boldsymbol{E}(\lambda + \epsilon)) = T(I)$ for $\lambda \geqq b$ and $\epsilon > 0$. But then $\boldsymbol{E}(\lambda)$ is the identity in $\mathfrak{L}$ for $\lambda \geqq b$. Thus we have (3).

Next we note that for $\epsilon_1 > \epsilon > 0$, we have

$$\boldsymbol{E}(\lambda) \leqq \boldsymbol{E}(\lambda + \epsilon) \leqq T(\boldsymbol{E}(\lambda + \epsilon_1)).$$

Hence, allowing ϵ and ϵ_1 both to approach zero, preserving the relations $\epsilon_1 > \epsilon$, $\epsilon_1 > 0$, we obtain (4).

Finally, we consider the behavior of $T(A)$ in the range of $\boldsymbol{E}(\lambda)$ and its orthog-

onal complement. Since $E(\lambda)AE(\lambda)$ has upper bound not exceeding λ, it follows that the upper bound of $T(E(\lambda+\epsilon))T(A)T(E(\lambda+\epsilon))$ does not exceed $\lambda+\epsilon$. But for every f in $\mathfrak{L}$, we have

$$(\boldsymbol{E}(\lambda)T(A)\boldsymbol{E}(\lambda)f, f) = \lim_{\epsilon\to 0}(T(E(\lambda+\epsilon))T(A)T(E(\lambda+\epsilon))f, f), \qquad \epsilon > 0.$$

Hence the upper bound of $\boldsymbol{E}(\lambda)T(A)\boldsymbol{E}(\lambda)$ does not exceed λ and (5) is established.

In entirely similar fashion, it can be shown that in the range of $1-\boldsymbol{E}(\lambda)$ the lower bound of $T(A)$ is not less than λ. Moreover, if λ is the lower bound and this lower bound is attained, we have for some f in $\mathfrak{L}$,

$$(1-\boldsymbol{E}(\lambda))T(A)(1-\boldsymbol{E}(\lambda))f = \lambda f.^{30}$$

Hence

$$\lim_{\epsilon\to 0}(T(A)(1-\boldsymbol{E}(\lambda+\epsilon))f,(1-\boldsymbol{E}(\lambda+\epsilon))f) = \lambda\,|f|^2.$$

But $(T(A)\ (1-\boldsymbol{E}(\lambda+\epsilon))f,\ (1-\boldsymbol{E}(\lambda+\epsilon))f)$ is monotone non-decreasing as ϵ approaches zero, and consequently we have

$$(T(A)(1-\boldsymbol{E}(\lambda+\epsilon))f,(1-\boldsymbol{E}(\lambda+\epsilon))f) \leqq \lambda\,|f|^2$$

for some $\epsilon > 0$. This, however is impossible, since in the range of $1-\boldsymbol{E}(\lambda+\epsilon)$, the lower bound of $T(A)$ is not less than λ. Hence, we have (6).

These six properties are sufficient to characterize $\boldsymbol{E}(\lambda)$ as the resolution of the identity of $T(A)$. One can for example, argue as follows: On the basis of these six properties the approximation theorem (Theorem 6) of the paper [6] of Lengyel and Stone can be proved and from this result it follows that $\boldsymbol{E}(\lambda)$ permutes with every bounded linear operator defined over $\mathfrak{L}$ which permutes with $T(A)$ ([6], Theorem 7). But this fact together with properties (1), (5), (6) above uniquely determines $\boldsymbol{E}(\lambda)$ as the resolution of the identity for $T(A)$ ([6], Theorem 5).

We are now in position to establish certain relationships between the resolutions of the identity belonging to two self-adjoint transformations in $\mathfrak{H}$ whose difference is totally continuous.

THEOREM 5.11. *Let A and B be self-adjoint transformations in $\mathfrak{H}$ such that $A-B$ is totally continuous. Let $E(\lambda)$ and $F(\lambda)$ be the resolutions of the identity corresponding to A and B respectively. Then, if μ is in the augumented resolvent set of A, $E(\mu)-F(\mu)$ is totally continuous.*

Let $\boldsymbol{E}(\lambda)$ be the resolution of the identity of $T(A)$ in $\mathfrak{L}$. Then, since μ is in the resolvent set of $T(A)$, we have for some $\delta > 0$, $\boldsymbol{E}(\mu-\delta) = \boldsymbol{E}(\mu+\delta)$. But, from Theorem 5.10 it is clear that we have

$$\boldsymbol{E}(\mu-\delta) \leqq T(E(\mu)) \leqq \boldsymbol{E}(\mu+\delta),$$

$$\boldsymbol{E}(\mu-\delta) \leqq T(F(\mu)) \leqq \boldsymbol{E}(\mu+\delta).$$

[30] [6], Theorem 3, for example.

Thus $T(E(\mu)) = T(F(\mu))$ and hence $E(\mu) - F(\mu)$ belongs to $\mathfrak{T}$, as we wished to show.

It is important to note that the requirement that μ be in the augmented resolvent set cannot be dropped from Theorem 5.11. Consider, for example, a nonnegative definite self-adjoint transformation A in $\mathfrak{H}$ which is totally continuous and has an inverse, and its resolution of the identity $E(\lambda)$. Consider also the transformation 0 in $\mathfrak{H}$, and its resolution of the identity $F(\lambda)$. Then $A - 0$ is in $\mathfrak{T}$, but $E(0) = 0$ and $F(0) = I$.

THEOREM 5.12.[31] *Let A and B be self-adjoint transformations in $\mathfrak{H}$ such that $A - B$ is totally continuous, and let $E(\lambda)$ and $F(\lambda)$ be their respective resolutions of the identity. Then, if μ is less than λ, there exists a totally continuous transformation $T_{\lambda,\mu}$, in B such that $E(\mu) + T_{\mu,\lambda}$ is a projection satisfying the inequality*

$$E(\mu) + T_{\mu,\lambda} \leqq F(\lambda).$$

If $T_{\mu,\lambda}$ can be chosen so that the equality sign holds, then the only points λ_0 on the interval $\mu < \lambda_0 < \lambda$ which belong to the spectrum of either A or B are characteristic values of finite multiplicity.

Conversely, let A and B be self-adjoint transformations in $\mathfrak{H}$ with resolutions of the identity $E(\lambda)$ and $F(\lambda)$, respectively. Then, if the inequalities $T(E(\mu)) \leqq T(F(\lambda))$ and $T(F(\mu)) \leqq T(E(\lambda))$ hold for all λ and μ such that $\mu < \lambda$, $A - B$ is totally continuous.

We prove the converse part first. From the inequalities in question, it follows at once that

$$\lim_{\epsilon \to 0} T(E(\lambda + \epsilon)) = \lim_{\epsilon \to 0} T(F(\lambda + \epsilon)), \qquad \epsilon > 0,$$

for all λ and thus that $T(A)$ and $T(B)$ have the same resolution of the identity. Hence $T(A) = T(B)$, or $A - B$ is in $\mathfrak{T}$.

Now let $A - B$ belong to $\mathfrak{T}$, and let $\boldsymbol{E}(\lambda)$ be the resolution of the identity of $T(A)$ in $\mathfrak{L}$. Then, for $\mu < \lambda$, we have

$$T(\boldsymbol{E}(\mu)) \leqq \boldsymbol{E}(\mu) \leqq T(F(\lambda)).$$

Hence $T(F(\lambda))T(\boldsymbol{E}(\mu))T(F(\lambda)) = T(\boldsymbol{E}(\mu))$ and therefore $F(\lambda)\boldsymbol{E}(\mu)F(\lambda) - \boldsymbol{E}(\mu)$ is in $\mathfrak{T}$. Consequently, invoking Theorem 2.4 with reference to the ring of bounded everywhere defined operators in the range of $F(\lambda)$, we see that there exists a totally continuous operator T_0 in $\mathfrak{B}$, with the value zero in the range of $I - F(\lambda)$, such that

$$F(\lambda)\boldsymbol{E}(\mu)F(\lambda) + T_0$$

is a projection satisfying

$$F(\lambda)\boldsymbol{E}(\mu)F(\lambda) + T_0 \leqq F(\lambda).$$

[31] Compare §4 of [20] to which this theorem is closely related.

Thus, if we set

$$T_{\mu,\lambda} = T_0 + F(\lambda)E(\mu)F(\lambda) - E(\mu),$$

$T_{\mu,\lambda}$, is in $\mathfrak{T}$ and $E(\mu) + T_{\mu,\lambda}$ is a projection satisfying the inequality of the theorem.

Now suppose we have

$$E_\mu + T_{\mu,\lambda} = F(\lambda).$$

Then $T(E(\mu)) = T(F(\lambda))$, and since for all λ_0 on the interval $\mu < \lambda_0 < \lambda$, we have

$$T(E(\mu)) \leqq \boldsymbol{E}(\lambda_0) \leqq T(F(\lambda)),$$

it follows that $\boldsymbol{E}(\lambda)$ is constant on that interval. Hence every point on the interval is in the resolvent set of $T(A) = T(B)$, and this is equivalent to the concluding statement of the first paragraph of the theorem.

COROLLARY. *Let A, B, $E(\lambda)$, and $F(\lambda)$ be as in Theorem* 5.12. *Then, if* $\lambda \neq \mu$,

$$E(\mu)F(\lambda) - F(\lambda)E(\mu)$$

is totally continuous.

From the inequality of Theorem 5.12, we have, for $\mu < \lambda$,

$$F(\lambda)E(\mu) + F(\lambda)T_{\mu,\lambda} = E(\mu)F(\lambda) + T_{\mu,\lambda}F(\lambda),$$

or

$$F(\lambda)E(\mu) - E(\mu)F(\lambda) = T_{\mu,\lambda}F(\lambda) - F(\lambda)T_{\mu,\lambda}.$$

Thus $F(\lambda)E(\mu) - E(\mu)F(\lambda)$ is totally continuous for $\mu < \lambda$, and by symmetry for $\mu > \lambda$ also.

This is not necessarily so, however, for $\lambda = \mu$. For consider any two projections E and F in $\mathfrak{H}$ such that $\mathfrak{R}(E)$ and $\mathfrak{R}(F)$ and their orthogonal complements are Hilbert spaces. Let A be a transformation which is equal to zero in $\mathfrak{R}(E)$ and which induces in $\mathfrak{H} \ominus \mathfrak{R}(E)$ a nonnegative definite self-adjoint totally continuous transformation whose inverse exists. Then A is in $\mathfrak{T}$ and if $E(\lambda)$ is the resolution of the identity of A, $E(0) = E$. Similarly, we can define a self-adjoint totally continuous operator B with resolution of the identity $F(\lambda)$ such that $F(0) = F$. But then $A - B$ is in $\mathfrak{T}$, while in general $E(0)F(0) - F(0)E(0)$ is not.

It follows also from the preceding example that the hypothesis $\mu < \lambda$ of Theorem 5.12 cannot be replaced by the hypothesis $\mu \leqq \lambda$.

We conclude this paper with a few observations concerning congruence modulo $\mathfrak{T}$ in $\mathfrak{B}$ which are not restricted to self-adjoint transformations.

THEOREM 5.13. *Let A_1 and A_2 be members of $\mathfrak{B}$ whose difference belongs to $\mathfrak{T}$, $A_1 = W_1B_1$, $A_2 = W_2B_2$ their canonical decompositions. Then $B_1 - B_2$ is in $\mathfrak{T}$, and if zero is in the augmented resolvent set of B_1, $W_1 - W_2$ is in $\mathfrak{T}$.*

Since $A_1^*A_1 = B_1^2$, $A_2^*A_2 = B_2^2$, we have $T(A_1^*A_1) = T(B_1^2) = T(B_2^2)$. More-

over, $[T(B_1)]^2 = [T(B_2)]^2 = T(B_1^2)$. But B_1 and B_2 are nonnegative definite and hence

$$\operatorname*{Lim}_{n\to\infty}(B_1 f_n, f_n) \geqq 0, \qquad \operatorname*{Lim}_{n\to\infty}(B_2 f_n, f_n) \geqq 0$$

for all $\{f_n\}$ in $\mathfrak{F}''$, by property (f) of $\operatorname*{Lim}_{n\to\infty}$. Hence $T(B_1)$ and $T(B_2)$ are both nonnegative definite, and since $T(B_1^2)$ can have only one nonnegative definite square root, we have $T(B_1) = T(B_2)$ which implies that $B_1 - B_2$ is in $\mathfrak{T}$.

Now let

$$T(A_1) = \boldsymbol{WB}$$

be the canonical decomposition of $T(A_1)$. Then $\boldsymbol{B} = T(B_1) = T(B_2)$. Moreover,

$$T(A_1) = T(W_1)T(B_1) = T(W_2)T(B_2).$$

Hence we can infer that the initial set of the partially isometric transformation $T(W_1)$ contains the initial set of $\boldsymbol{W}$ (the closure of the range of $\boldsymbol{B}$) and that the two transformations are equal there. Hence $T(W_1)$ is identical with $\boldsymbol{W}$ provided the initial set of $T(W_1)$ is the closure of the range of $\boldsymbol{B}$. But, if the origin is in the augmented resolvent set of B_1, it is in the resolvent set of $T(B_1) = \boldsymbol{B}$ and hence $\boldsymbol{W}$ is unitary. Thus under the hypothesis of the theorem we have $T(W_1) = \boldsymbol{W}$. Furthermore, under that hypothesis, the origin is also in the augmented resolvent set of B_2 and hence a similar argument yields the equation $T(W_2) = \boldsymbol{W}$. Thus $T(W_1) = T(W_2)$ and $W_1 - W_2$ is in $\mathfrak{T}$ as we wished to show.

We wish to emphasize that the hypothesis that zero be in the augmented resolvent set of B_1 cannot be omitted from Theorem 5.13. For example, consider a totally continuous nonnegative definite self-adjoint transformation A with an inverse. Then if $A = W_1B_1$ is its canonical decomposition and $0 = W_2B_2$ is the canonical decomposition of 0, we have $W_1 = I$, $W_2 = 0$, $A - 0$ in $\mathfrak{T}$.

We now extend Theorem 5.9 to cover transformations which are not necessarily self-adjoint.

DEFINITION 5.3. *Let A be an arbitrary bounded everywhere defined transformation in $\mathfrak{H}$. We define the augmented resolvent set of A as the set of points λ in the complex plane such that the following conditions are satisfied:*

(1) *the manifold of zeros of $A - \lambda I$ has a finite dimension number;*

(2) *the range of $A - \lambda I$ is closed;*

(3) *the orthogonal complement of the range of $A - \lambda I$ has a finite dimension number.*

The complement of the augmented resolvent set is called the condensed spectrum of A.

Evidently the augmented resolvent set of A contains the resolvent set, so that our terminology is justified.

THEOREM 5.14. *Let A and C be two bounded everywhere-defined transforma-*

tions in $\mathfrak{H}$ such that $A - C$ is totally continuous. Then A and C have the same augmented resolvent sets.

It is sufficient for the proof to show that every point in the augmented resolvent set of A is also in the augmented resolvent set of C.

Let λ be a point of the augmented resolvent set of A, $\mathfrak{M}$ the manifold of zeros of $A - \lambda I$. Then on $\mathfrak{H} \ominus \mathfrak{M}$, $A - \lambda I$ induces a transformation T with an inverse whose range is $\mathfrak{R}(A - \lambda I)$. Moreover, since $\mathfrak{R}(A - \lambda I)$ is closed, T^{-1} is bounded.

Now let $\mathfrak{N}$ be the manifold of zeros of $C - \lambda I$. Then if $\mathfrak{N}$ is a Hilbert space, $(\mathfrak{H} \ominus \mathfrak{M}) \cdot \mathfrak{N}$ is a Hilbert space, and T has a contraction T_1 with domain $(\mathfrak{H} \ominus \mathfrak{M}) \cdot \mathfrak{N}$ which also has a bounded inverse. But in $(\mathfrak{H} \ominus \mathfrak{M}) \cdot \mathfrak{N}$, we have $A - C = T_1$ and this contradicts our hypothesis that $A - C$ is in $\mathfrak{T}$. Hence $\mathfrak{N}$ must have a finite dimension number.

We now observe that when λ is a point of the augmented resolvent set of A, $\bar{\lambda}$ is a point of the augmented resolvent set of A^*. Hence the preceding argument serves to show that the manifold of zeros of $C^* - \lambda I$ has a finite dimension number. But this manifold is precisely the orthogonal complement of the closure of $\mathfrak{R}(C - \lambda I)$, and hence to show that λ is in the augmented resolvent set of C, it remains only for us to prove that $\mathfrak{R}(C - \lambda I)$ is closed.

To establish the latter we consider the canonical decompositions

$$A - \lambda I = W_1 B_1 , \qquad C - \lambda I = W_2 B_2 .$$

From Theorem 5.13, it follows that $B_1 - B_2$ is totally continuous. Moreover, since the range of $A - \lambda I$ is closed, so also is the range of B_1, while the manifold of zeros of B_1 is the manifold of zeros of $A - \lambda I$. But then the origin is in the augmented resolvent set of B_1 and hence in the augmented resolvent set of B_2. Hence $\mathfrak{R}(B_2)$ is closed which implies that $\mathfrak{R}(C - \lambda I)$ is closed also as we wished to show.

If λ is in the augmented resolvent set of A, we denote by R_λ the transformation which is equal to zero in $\mathfrak{H} \ominus \mathfrak{R}(A - \lambda I)$ and which takes each element f of $\mathfrak{R}(A - \lambda I)$ into that element g in the orthogonal complement of the manifold of zeros of $(A - \lambda I)$ which satisfies $(A - \lambda I)g = f$. Thus $R_\lambda(A - \lambda I)$ is the projection with range the orthogonal complement of the manifold of zeros of $A - \lambda I$. We call the family of transformations R_λ so defined the augmented resolvent of A.

THEOREM 5.15. *Let A and B be two transformations in $\mathfrak{H}$ such that $A - B$ is in $\mathfrak{T}$. Let $R_\lambda^{(1)}$ and $R_\lambda^{(2)}$ be respectively their augmented resolvents. Then $R_\lambda^{(1)} - R_\lambda^{(2)}$ is in $\mathfrak{T}$ for all λ for which these transformations are defined.*

Consider the transformation $T(A) = T(B)$ in $\mathfrak{L}$. If λ is in the augmented resolvent set of A (or B) it is in the resolvent set of $T(A)$, and $T(R_\lambda^{(1)})$ and $T(R_\lambda^{(2)})$ are both inverses of $T(A) - \lambda \cdot 1 = T(B) - \lambda \cdot 1$. But only one such inverse can exist; hence $T(R_\lambda^{(1)}) = T(R_\lambda^{(2)})$ and $R_\lambda^{(1)} - R_\lambda^{(2)}$ is in $\mathfrak{T}$.

ILLINOIS INSTITUTE OF TECHNOLOGY

REFERENCES

1. A. A. Albert, *Modern Higher Algebra*, Chicago, 1937.
2. S. Banach, *Théorie des opérations linéaires*, Warsaw, 1932.
3. J. W. Calkin, *Abstract Symmetric Boundary Conditions*, Transactions of the American Mathematical Society, 45 (1939), pp. 369–442.
4. R. Courant, *Zur Theorie der linearen Integralgleichungen*, Mathematische Annalen, 89 (1923), pp. 161–178.
5. K. Friedrichs, *Spektraltheorie halbbeschrankten Operatoren*, Mathematischen Annalen, 109 (1934), pp. 465–487.
6. B. A. Lengyel and M. H. Stone, *Elementary Proof of the Spectral Theorem*, Annals of Mathematics, 37 (1936), pp. 853–864.
7. H. Löwig, *Komplexe euclidische Räume von beliebiger endlicher oder transfiniter Dimensionzahl*, Acta Litterarum ac Scientiarum, Sectio Scientiarum Mathematicarum, Szeged, 7 (1934), pp. 1–33.
8. F. J. Murray, *Linear Transformations between Hilbert Spaces and the Applications of this Theory to Linear Partial Differential Equations*, Transactions of the American Mathematical Society, 37 (1935), pp. 301–338.
9. F. J. Murray and J. von Neumann, *On Rings of Operators*, I, Annals of Mathematics, 37 (1936), pp. 116–229.
10. J. von Neumann, *Allgemeine Eigenwerttheorie Hermitescher Funktionaloperatoren*, Mathematische Annalen, 102 (1929), pp. 49–131.
11. J. von Neumann, *Zur Algebra der Funktionaloperatoren*, ibid., 102 (1929), pp. 370–427.
12. J. von Neumann, *Uber adjungtierte Funktionaloperatoren*, Annals of Mathematics, 33 (1932), pp. 294–310.
13. J. von Neumann, *Lectures on Operator Theory* (mimeographed), Princeton, 1934.
14. J. von Neumann, *Charakterisierung des Spektrums eines Integraloperators*, Actualités Scientifiques et Industrielles 229, Paris, 1935.
15. J. von Neumann, *On Rings of Operators*, III, Annals of Mathematics, 41 (1940), pp. 94–161.
16. F. Rellich, *Spektraltheorie in nichtseparablen Räumen*, Mathematische Annalen, 110 (1935), pp. 342–356.
17. M. H. Stone, *Linear Transformations in Hilbert Space and their Applications to Analysis*, New York, 1932.
18. A. E. Taylor, *The Hilbert Space Postulates—A Further Reduction*, Bulletin of the American Mathematical Society, 41 (1935), pp. 847–848.
19. B. L. van der Waerden, *Moderne Algebra*, I, Berlin, 1930.
20. H. Weyl, *Uber beschränkte quadratische Formen, deren Differenz vollstetig ist*, Rendiconti del Circolo Matematico Palermo, 27 (1909), pp. 373–392.

ANNALS OF MATHEMATICS
Vol. 42, No. 4, October, 1941

MEASURE-PRESERVING HOMEOMORPHISMS AND METRICAL TRANSITIVITY*

BY J. C. OXTOBY AND S. M. ULAM

(Received May 1, 1941)

INTRODUCTION

In the study of dynamical systems one is led naturally to the consideration of measure-preserving transformations. A Hamiltonian system of $2n$ differential equations induces in the phase space of the system a measure-preserving flow, that is, a one-parameter group of transformations that leave invariant the $2n$-dimensional measure. Making use of one or more integrals of the system, one obtains a reduced phase manifold of lower dimension which likewise undergoes a flow into itself, and in general admits an invariant measure related in a simple manner to that in the larger phase space. If the differential equations are sufficiently regular the flow will have corresponding properties of continuity and differentiability. Thus the study of one-parameter continuous groups of measure-preserving automorphisms[1] of finite dimensional spaces has an immediate bearing on dynamics and the theory of differential equations.

In statistical mechanics one is especially interested in time-average properties of a system. In the classical theory the assumption was made that the average time spent in any region of phase space is proportional to the volume of the region in terms of the invariant measure, more generally, that time-averages may be replaced by space-averages. To justify this interchange, a number of hypotheses were proposed, variously known as ergodic or quasi-ergodic hypotheses, but a rigorous discussion of the precise conditions under which the interchange is permissible was only made possible in 1931 by the ergodic theorem of Birkhoff.[2] This established the *existence* of the time-averages in question, for almost all initial conditions, and showed that if we neglect sets of measure zero, the interchange of time- and space-averages is permissible if and only if the flow in the phase space is *metrically transitive*. A transformation or a flow is said to be metrically transitive if there do not exist two disjoint invariant sets both having

---

* This paper includes results presented to the Amer. Math. Soc. in preliminary reports on April 15 and December 30, 1938. Most of the results were obtained while the authors were members of the Harvard Society of Fellows.

[1] An automorphism is a 1:1 bicontinuous transformation of a space onto itself. It is measure-preserving with respect to a measure μ if $\mu TA = \mu A = \mu T^{-1}A$ whenever A is measurable.

[2] G. D. Birkhoff, *Proof of the ergodic theorem*. Proc. Nat. Acad. USA. **17** (1931) 650–660. An interesting connection between this theorem and the fundamental theorem of the calculus has been shown by N. Wiener. See *The ergodic theorem*. Duke Jour. **5** (1939) 1–18.

positive measure.[3] Thus the effect of the ergodic theorem was to replace the ergodic hypothesis by the hypothesis of metrical transitivity.[4]

Nevertheless, in spite of the simplification introduced by the ergodic theorem, the problem of deciding whether particular systems are metrically transitive or not has proved to be very difficult. Hedlund[5] showed that the flow defined by the system of geodesics on certain surfaces of constant negative curvature is metrically transitive, and this result was generalized by Hopf.[6] More recently, the result has been extended to surfaces of variable negative curvature, independently by both Hedlund[7] and Hopf.[8] This important class of systems is so far the only large class known to be metrically transitive. In fact, the only other known examples of metrically transitive continuous flows are the spiral motions on the n-dimensional torus, including the rotation of the circumference of a circle as the simplest example of all. The ergodic properties of these systems were established as long ago as 1916 by Weyl,[9] in a quite different connection.

Thus the known examples of metrically transitive continuous flows are all in manifolds, indeed in manifolds of restricted topological type, either toruses or manifolds of direction elements over surfaces of negative curvature.[10] An outstanding problem in ergodic theory has been the existence question—can a metrically transitive continuous flow exist in an arbitrary manifold, or in any space that is not a manifold? In the present paper we shall obtain a complete answer to this question, at least on the topological level, for polyhedra of dimension three or more. It will appear that the only condition that needs to be imposed is a trivially necessary kind of connectedness. In particular, there exists a metrically transitive continuous flow in the cube, in the solid torus, and in any pseudo-manifold of dimension at least three. Since the phase spaces of dynamical systems have the required kind of connectedness, it follows that the hypothesis of metrical transitivity in dynamics involves no *topological* contradiction. More precisely, in any phase space there can exist a continuous flow

[3] The notion was first introduced in a different connection by G. D. Birkhoff and P. A. Smith. See *Structure analysis of surface transformations*. Liouville's Jour. **7** (1928) 365.

[4] For a historical survey, see G. D. Birkhoff and B. O. Koopman, *Recent contributions to the ergodic theory*. Proc. Nat. Acad. USA. **18** (1932) 279.

[5] G. A. Hedlund, *On the metrical transitivity of the geodesics on closed surfaces of constant negative curvature*. Ann. of Math. **35** (1935) 787. See also his paper, *A new proof for a metrically transitive system*. Amer. J. Math. **62** (1940) 233–242.

[6] E. Hopf, *Fuchsian groups and ergodic theory*. Trans. Amer. Math. Soc. **39** (1936) 299. See also his booklet, *Ergodentheorie*. Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 5, Berlin 1937.

[7] G. A. Hedlund, Bull. Amer. Math. Soc. Abstract 46-3-173.

[8] E. Hopf, *Statistik der geodätischen Linien in Mannigfaltigkeiten negativer Krümmung*. Ber. Verh. Sächs. Akad. Wiss. Leipzig **91** (1939) 261–304.

[9] H. Weyl, *Über die Gleichverteilung von Zahlen mod. Eins*. Math. Ann. **77** (1916) 315.

[10] Some examples of discontinuous flows have been studied. See E. Hopf, *Ergodentheorie*, and J. von Neumann, *Zur Operatorenmethode in der klassischen Mechanik*. Annals of Math. **33** (1932) 587–642.

metrically transitive with respect to the invariant measure associated with the system.

It may be recalled that the original ergodic hypothesis of Boltzmann—that a single streamline passes through all points of phase space—had to be abandoned because it involved a topological impossibility.[10a] It was replaced by a quasi-ergodic hypothesis—that some streamline passes arbitrarily close to all points of phase space. But it is not obvious that even this weak hypothesis is topologically reasonable in general phase spaces, and in any case it is not sufficient to justify the interchange of time- and space-averages. It is therefore of some interest to know that the ergodic hypothesis in its modern form of metrical transitivity is at least free from any objection on topological grounds.

It must be emphasized, however, that our investigation is on the topological level. The flows we construct are continuous groups of measure-preserving automorphisms, but not necessarily differentiable or derivable from differential equations. Thus they correspond to dynamical systems only in a generalized sense. In this respect the flows studied by Hedlund and Hopf are closer to the dynamical problem. In any case, the problem is illuminated by considering it in a more general setting, and probably a certain amount of generalization is necessary in order to give meaning to some questions of the type with which we deal.

The conjecture has frequently been expressed, first by Birkhoff, then by Hopf and others, that metrical transitivity is probably the "general case." The conjecture was never given a precise formulation, but was based on the fact that the transitive case is the non-integrable case, in the sense that no uniform measurable integrals exist, and on the idea that a general transformation or flow should be expected to shuffle points in a more or less random fashion, and therefore leave invariant as few sets as possible. To see the precise connection between metrical transitivity and randomness, let us recall the statement of the ergodic theorem for transformations. It asserts that if μ is a completely additive finite measure in E, and T a measure-preserving transformation of E onto itself, then for any μ-integrable function $f(p)$ on E the average value $\lim_{n\to\infty} \frac{1}{n} \sum_{r=1}^{n} f(T^r p) = f^*(p)$ exists for almost all points p, is a measurable function of p, and that $\int f^*(p)\, d\mu(p) = \int f(p)\, d\mu(p)$. Metrical transitivity is precisely the condition that $f^*(p)$ should be constant for almost all p. Thus if T is metrically transitive and f_A is the characteristic function of any measurable set A, we have $\frac{\mu A}{\mu E} = \lim_{n\to\infty} \frac{1}{n} \sum_{r=1}^{n} f_A(T^r p)$ for almost all p. The left member of this equation represents the average number of images of p that fall in A, or the *frequency* with which the images of p fall in the set A under iteration of T, and the equation

[10a] A. Rosenthal, *Beweis der Unmöglichkeit ergodischer Gassysteme*, Ann. der Physik, (4) **42** (1913) 796–806.

asserts that this frequency is proportional to the measure of A. A sequence of points selected at random from E would be expected to have just such a distribution, in fact, the ratio $\mu A/\mu E$ may be interpreted as the probability that a point selected at random lies in A. Thus a metrically transitive transformation is one under which almost all points generate sequences that are distributed like random sequences in respect to the average number that fall in any measurable set. This perhaps makes it seem plausible that a measure-preserving transformation "selected at random" should be metrically transitive.

One might try to make precise the idea that metrical transitivity is the general case by introducing a *measure* in the space of all measure-preserving transformations of E, but it seems difficult to do this in any natural way. Nevertheless, there is a simple and natural *metric* in the space of automorphisms of E, in case E is compact, and one may ask whether in this space metrical transitivity is the general case in the *topological sense*, that is, whether such automorphisms constitute all but a set of first category.[11] This is what we are going to show, under suitable assumptions about E. Thereby we shall dispose of the existence problem at the same time.[12] The result is perhaps a little surprising in view of the fact that metrical transitivity is a purely measure-theoretic property, and it often happens that what is the general case in the sense of measure is exceptional in the sense of category, and vice versa.[13]

The fact that metrical transitivity turns out to be the general case in the topological sense raises the question how far residual sets of measure-preserving automorphisms may exhibit other random properties. For instance, a random sequence is not only distributed so that the average number of points in a set is proportional to its measure, but the limiting frequency is approached in the manner of "Laplace-Liapounoff." Presumably metrical transitivity is not sufficient to insure that almost all points should generate sequences exhibiting this more precise behavior. It would therefore be of interest to know whether transformations of this sort are likewise general, or indeed whether they can exist

---

[11] A set is said to be of first category if it can be represented as a sum of countably many nowhere dense sets. Any other set is said to be of second category. Complements of first category sets are called residual sets.

[12] Baire's Theorem asserts that in a complete metric space every residual set is of second category and is dense. It summarizes concisely a typical form of existence proof.

[13] For instance, the law of large numbers is false in the sense of category. That is, the set of numbers x in $0 < x < 1$ such that in their infinite dyadic development the number of ones in the first n places divided by n tends to one-half is of first category (although of measure one). If x_n denoted the n-th digit in the dyadic development of x, the set in question is represented by

$$\prod_k \sum_N \left[\prod_n \mathsf{E}_x \left\{ \left| \frac{x_1 + x_2 + \cdots + x_{N+n}}{N+n} - \frac{1}{2} \right| < \frac{1}{k} \right\} \right],$$

and for $N > 0$ and $k > 2$ the set enclosed in square brackets is nowhere dense, as may be seen by inserting a sufficiently long block of zeros far out in the development of a number.

at all. A result in this direction is given in §12. In the same order of ideas one may ask for automorphisms exhibiting the various types of *mixture* properties.[14]

In the course of our investigation we shall derive a number of results about transformations and measures which have an interest entirely apart from questions relating to transitivity. In particular, the results obtained in Part II concerning measures topologically equivalent to Lebesgue measure appear to be fundamental to the theory of measure-preserving homeomorphisms. In sections 13 to 15 a number of questions of group-theoretic interest are answered on the basis of these results. The paper involves an intimate combination of the methods of topology and measure theory. A suggestive example of the way in which such a combination of methods may lead to purely topological results not otherwise apparent is the sort of topological ergodic theorem obtained in §16. Another result of purely topological character is the Corollary to Lemma 16 concerning equivalence of Cantor sets under automorphism of the containing space. This result is of independent interest since it adds to the results obtained by Antoine[15] relating to this question.

A fundamental outstanding problem in topology is that of approximating to an arbitrary homeomorphism by a differentiable one.[16] In §18, as a by-product of our investigation, we obtain the result that any measure-preserving automorphism of the r-dimensional cube that leaves the boundary fixed can be approximated uniformly by one that is differentiable *almost everywhere*, in fact, it is locally linear about almost all points. This result in itself is not strong enough to have important topological implications, but it suggests that the special properties of measure-preserving automorphisms may enable one to obtain approximation theorems for them that are difficult or perhaps impossible to establish in general. In any case, the very precise properties of the approximating transformation obtained in Theorem 12 may well serve as a basis for answering other questions concerning approximation by automorphisms with special measure-theoretic properties.

I. Preliminary Results

1. Definitions and Principal Theorem

We assume the standard notions[17] of *polyhedron*, *euclidean polyhedron*, and *complex*, except that we shall always understand these to be finite. Thus a polyhedron is compact. Some of the results can be generalized to infinite polyhedra, but it seems best to leave such extensions out of the present paper.

---

[14] See G. A. Hedlund, *The dynamics of geodesic flows*. Bull. Amer. Math. Soc. **45** (1939) 241–260.

[15] L. Antoine, *Sur l'homéomorphie de deux figures et de leurs voisinage*. Liouville's Jour. (8) **4** (1921) 221–325 esp. p. 307 ff.

[16] J. W. Alexander, *Some problems in topology*. Verh. des Int. Math. Kong. Zurich 1932, vol. 1, p. 249–257.

[17] For all topological definitions see P. Alexandroff and H. Hopf, *Topologie I*, Berlin 1935, hereafter referred to as AH.

A *regular point* of an r-dimensional polyhedron is a point that has a neighborhood homeomorphic to an open sphere in r-space, any other point of the polyhedron is called *singular*. A polyhedron is called *regularly connected* if its regular points form a connected set dense in the polyhedron.[18]

A *finite outer measure* in a space E is a function μ^* defined for all subsets of E and satisfying the three conditions:

M1: $0 \leqq \mu^* A \leqq \mu^* E, \quad o < \mu^* E < +\infty, \quad \mu^*(\text{void set}) = 0$

M2: $\mu^* A \leqq \mu^* B \quad$ if $A \subset B$

M3: $\mu^* \sum_{n=1}^{\infty} A_n \leqq \sum_{n=1}^{\infty} \mu^* A_n$.

A set A is *measurable* with respect to μ^* if $\mu^* W = \mu^* WA + \mu^* W(E - A)$ for every set W. The function μ^* is completely additive with respect to its class of measurable sets, and the measure function thus derived from μ^* by restricting its domain is denoted by μ.[19]

A *Carathéodory outer measure* is one that satisfies also the condition

M4: $\mu^*(A + B) = \mu^* A + \mu^* B$ if A and B are separated by a positive distance.

The measurable sets then include all Borel sets.

By a *Lebesgue-Stieltjes outer measure* in a polyhedron, we shall understand a finite Carathéodory outer measure that satisfies the further condition

M5: $\mu^* A = \inf_{G \supset A} \mu^* G \quad G$ open.

A measure derived from such an outer measure will be called a Lebesgue-Stieltjes measure (LS measure).[20]

We introduce the following special definition. A Lebesgue-Stieltjes measure in an r-dimensional polyhedron E, $r \geqq 1$, will be called *r-dimensional* if it is zero for points, zero for the set of singular points, and positive for neighborhoods of regular points. This definition is consistent with ordinary usage, since r-dimensional Lebesgue measure in any r-dimensional euclidean polyhedron evidently fulfills the requirements. The invariant measures associated with dynamical systems are usually defined by integrating a positive density function and are therefore r-dimensional in the present sense.

The set of all automorphisms of a polyhedron E (or of any compact space) is made into a metric space $H[E]$ by the definition

$$\rho(g, h) = \max \left[\max_{x \in E} \rho_E(gx, hx), \max_{x \in E} \rho_E(g^{-1}x, h^{-1}x) \right],$$

[18] This is equivalent to the ordinary definition. See AH pp. 400, 402.

[19] For the theory of measure, see Carathéodory, *Vorlesungen über reellen Funktionen*, Leipzig and Berlin 1918; or H. Hahn, *Theorie der reelen Funktionen I*, Berlin 1921.

[20] For polyhedra embedded in euclidean spaces, these measures are the same as those obtained by relativizing measures in the containing space that are derived from non-negative additive functions of an interval. The reasoning is essentially contained in S. Saks, *Theory of the integral*, Warsaw-Lwów 1937, Chap. 3. Thus our definition is consistent with the ordinary concept of a non-negative Lebesgue-Stieltjes measure, and has the advantage of being intrinsic. Also it is equally applicable to curved polyhedra, in which intervals have no meaning. An outer measure that satisfies conditions M1 to M5 is called by Hahn a (finite) Inhaltsfunktion. See Hahn, op. cit. p. 444.

where g, h are any two automorphisms of E, and ρ_E denotes distance in E. In this metric, convergence of a sequence of automorphisms means uniform convergence together with uniform convergence of the sequence of inverses. One can verify without difficulty that the space $H[E]$ is complete and that the group operations (composition and inverse of automorphisms) are continuous in this metric, so that $H[E]$ is a *metric group*, that is, a space of type (G) in the sense of (e.g.) Banach.[21] The distance of an automorphism from the identity will be called its *norm*, it is equal to the maximum distance through which any point is displaced by the automorphism. An *arbitrarily small* automorphism is one whose norm is arbitrarily small.

It may be remarked that whereas uniform convergence of a sequence of automorphisms h_n does not insure the existence of a limiting automorphism, nevertheless uniform convergence of h_n *to an automorphism* h implies uniform convergence of h_n^{-1} to h^{-1}, therefore convergence in $H[E]$.[22] Hence the metric in $H[E]$ is topologically equivalent to the metric $\max_{x \in E} \rho_E(gx, hx)$ usually used to metrize the space of continuous mappings of E into itself, but with respect to the latter metric the space of automorphisms is not complete.

The main object of our investigation, however, will not be the space $H[E]$ but rather the subspace consisting of all measure-preserving automorphisms with respect to a given LS measure μ.[23] This subspace, with the same metric as in $H[E]$, will be denoted by $M[E, \mu]$. To see that $M[E, \mu]$ is a closed subset of $H[E]$, and therefore complete, consider any sequence of measure-preserving automorphisms T_n converging to the limit T in $H[E]$. Let F be any closed subset of E, and G any open set containing TF. From the uniform convergence of the sequence T_n it follows that $T_nF \subset G$, for all sufficiently large n, so that $\mu^*G \geqq \mu^*T_nF = \mu^*F$. Therefore $\mu^*F \leqq \mu^*TF$, by condition M5. Similar reasoning applied to the sequence T_n^{-1} yields the inverse inequality, so that $\mu^*TF = \mu^*F$ for all closed sets F. Hence equality holds also for all open sets, therefore for all sets, and the limiting automorphism T is therefore measure-preserving. Since the measure-preserving automorphisms of E form a group, it follows that $M[E, \mu]$ is a metric group, in fact, a closed subgroup of $H[E]$.

In case E is a rectangular r-cell R, and μ is the ordinary r-dimensional Lebesgue measure m in R, we shall write simply $M[R]$ instead of $M[R, m]$. The closed subgroups of $M[R]$ and $H[R]$ consisting of automorphisms that leave all boundary points fixed, we shall denote by $M_0[R]$ and $H_0[R]$ respectively.

THEOREM 1: *Let E be any regularly connected polyhedron of dimension $r \geqq 2$,*

[21] S. Banach, *Théorie des opérations linéaires*, Warsaw 1932, Chap. 1 and p. 229. Our metric in $H[E]$ is topologically equivalent to the one there assigned to this group.

[22] J. Schreier and S. Ulam, *Über topologische Abbildungen der euklidischen Sphären*. Fund. Math. **23** (1934) 102–118, esp. p. 104.

[23] We shall denote general measures by μ or ν. The letter m will be reserved for Lebesgue measure. General automorphisms will be denoted by g or h, measure-preserving automorphisms by S or T. For LS measures, the definition of a measure-preserving automorphism may be taken to be the condition $\mu^*TA = \mu^*A$ for every A.

and let μ be any r-dimensional Lebesgue-Stieltjes measure in E. In the space $M[E, \mu]$ of measure-preserving automorphisms of E the metrically transitive automorphisms form a residual G_δ set.[24]

The proof of this theorem will only be completed in Part III. In Part IV we shall discuss the generality of the result and its significance for metrically transitive flows.

In a previous note[25] one of us has obtained a similar result for the set of *topologically transitive* automorphisms, for slightly different assumptions on E. A transformation is topologically transitive if there do not exist two disjoint invariant open sets, both non-void, (or, equivalently, if the sequence of images of some point is dense in E). Metrical transitivity with respect to an r-dimensional LS measure obviously implies topological transitivity. Hence Theorem 1 may be considered as a generalization of the earlier result, but whereas the latter required only the most elementary constructions and properties of measure-preserving automorphisms, the present theorem will require a much more extensive investigation.

As regards the interpretation of Theorem 1 as a proof of the conjecture that metrical transitivity is the general case, it is interesting to note that the topological notion of probability in the sense of category can be subsumed under the general theory of measure and probability. In any complete metric space, if we define $\mu^*A = 0$ or 1 according as A is of first or second category, it is easily verified that conditions M1, M2, M3 are satisfied,and that the class of measurable sets consists of the first category sets and the residual sets, which have measures zero and one respectively. The notion of probability in the sense of category is therefore a special case of the general notion of (non-Borel) measure.

2. Preliminary Lemmas

The object of this first sequence of lemmas is to show that the general situation can be reduced, in all essential respects, to the consideration of ordinary Lebesgue measure in a rectangular r-cell. In making this reduction, a central role is played by the notion of a continuous map f of a convex cell Z onto a polyhedron E which is a *homeomorphism up to the boundary*. By this we shall mean that f maps the interior Z_0 of Z homeomorphically onto fZ_0. For such continuous maps $fBdZ = E - fZ_0$, so that $f^{-1}f$ is single-valued on Z_0, though possibly multiple-valued on BdZ.

LEMMA 1: *Let E be the polyhedron of a regularly connected euclidean complex K of dimension $r \geqq 1$. It is possible to represent E as the continuous image of a convex r-cell Z under a map f which is a homeomorphism up to the boundary and which is a simplicial map of a certain subdivision of Z onto K.*

This is not a new result, but rather a corollary of the known result that any

[24] A G_δ set is one that can be represented as a countable intersection of open sets.

[25] J. C. Oxtoby, *Note on transitive transformations*. Proc. Nat. Acad. USA. **23** (1937) 443–446. See also E. Hopf, *Statistische Probleme und Ergebnisse in der klassischen Mechanik*. Actualités Scientifiques et Industrielles, No. 737 (1938) 5–16.

regularly connected complex can be obtained from a convex cell by suitable boundary identifications.[26] To prove it, let $\sigma_1, \cdots, \sigma_n$ be the base simplexes of K, so numbered that each, after the first, has at least one regular face $\sigma_{i-1,i}$ in common with one of its predecessors. The possibility of such a numbering follows at once from the definition of a regularly connected complex.[27] Let τ_1 be any simplex in r-space, and set its vertices in correspondence with those of σ_1. Let $\tau_{1,2}$ be the face of τ_1 that corresponds to $\sigma_{1,2}$. Take a new vertex outside τ_1 which forms with $\tau_{1,2}$ a simplex τ_2 such that $\tau_1 + \tau_2$ is still convex. Any point sufficiently near the center of $\tau_{1,2}$ will suffice. Let this vertex correspond to the remaining vertex of σ_2. Evidently we have a simplicial map of the complex $\tau_1 + \tau_2$ onto the complex $\sigma_1 + \sigma_2$, and the only $(r-1)$-simplex in $\tau_1 + \tau_2$ which does not lie on the boundary is $\tau_{1,2}$. Proceeding by induction, suppose that we have a complex K_i with r-dimensional base simplexes $\tau_1, \cdots, \tau_i$ whose union is a convex cell, and a simplicial map onto the complex $\sigma_1 + \sigma_2 + \cdots + \sigma_i$ under which τ_j corresponds to σ_j, $1 \leqq j \leqq i$. Also suppose that the only $(r-1)$-simplexes of K_i that do not lie on the boundary are $\tau_{1,2}, \cdots, \tau_{i-1,i}$ which correspond to $\sigma_{1,2}, \cdots, \sigma_{i-1,i}$. The simplex $\sigma_{i,i+1}$ is a face of some one of $\sigma_1, \cdots, \sigma_i$. Let $\tau_{i,i+1}$ be the corresponding face of the corresponding one of the simplexes $\tau_1, \cdots, \tau_i$. Since $\sigma_{i,i+1}$ is regular, it is distinct from $\sigma_{1,2}, \cdots, \sigma_{i-1,i}$, and it follows that $\tau_{i,i+1}$ lies on the boundary of the convex cell. We can therefore adjoin a new simplex τ_{i+1} having $\tau_{i,i+1}$ for a face, and such that $\tau_1 + \cdots + \tau_{i+1}$ is again convex. If we map the new vertex of τ_{i+1} onto the remaining vertex of σ_{i+1} the simplicial map will be extended to map $K_{i+1} = \tau_1 + \cdots + \tau_{i+1}$ onto $\sigma_1 + \cdots + \sigma_{i+1}$, and the only $(r-1)$-dimensional simplexes of K_{i+1} not on its boundary will be $\tau_{1,2}, \cdots, \tau_{i,i+1}$. At the n-th stage we obtain a simplicial map of a complex K_n onto K. If we map the simplexes $\tau_1, \cdots, \tau_n$ affinely onto $\sigma_1, \cdots, \sigma_n$, we obtain a continuous map of a convex cell Z onto E. The interior of Z is the union of the interiors of $\tau_1, \cdots, \tau_n$ and of $\tau_{1,2}, \cdots, \tau_{n-1,n}$. These correspond to the interiors of $\sigma_1, \cdots, \sigma_n$ and of $\sigma_{1,2}, \cdots, \sigma_{n-1,n}$ respectively, which are disjoint since these simplexes are all distinct. Hence distinct interior points of Z go into distinct points, and the map is therefore a homeomorphism up to the boundary.

LEMMA 2: *Let f be a continuous map of a rectangular r-cell R, $r \geqq 1$, onto a polyhedron E which is a homeomorphism up to the boundary. Let μ be an r-dimensional Lebesgue-Stieltjes measure in E, and suppose that $\mu fBdR = 0$. Then the function $\nu^*A = \mu^* fA$ defines an r-dimensional Lebesgue-Stieltjes measure in R.*

That ν^* satisfies conditions M1 to M3 may be verified by inspection. Con-

---

[26] See AH p. 264.

[27] A complex is *homogeneous r-dimensional* if every simplex lies on an r-simplex. A *regular face* is an $(r-1)$-simplex that lies on two and only two r-simplexes of the complex. A homogeneous r-dimensional complex K is *regularly connected* if in any division of K into two r-dimensional subcomplexes, these have in common at least one regular face. It can be shown that a polyhedron is regularly connected if and only if it is the polyhedron of a regularly connected complex. See AH p. 402.

dition M4 is less evident in view of the fact that f need not be 1:1. To verify it, consider any two sets A, B contained in R and separated by a positive distance. Their closures $\bar{A}$, $\bar{B}$ are disjoint, and therefore $f\bar{A}$ and $f\bar{B}$ intersect at most in a subset of $fBdR$, since all other points have unique antecedents. Hence $\mu f\bar{A} \cdot f\bar{B} = 0$, by hypothesis. It is possible to enclose $f(A + B)$ in a G_δ set C such that $\mu^* f(A + B) = \mu C$. Let $A_1 = C \cdot f\bar{A}$ and $B_1 = C \cdot f\bar{B}$. Then we have $fA \subset A_1$ and $fB \subset B_1$, and so $\mu^* f(A + B) = \mu^*(fA + fB) \leqq \mu^*(A_1 + B_1) \leqq \mu^* C = \mu^* f(A + B)$, that is, $\mu^* f(A + B) = \mu(A_1 + B_1)$. But A_1 and B_1 are Borel sets that intersect in a set of measure zero, therefore $\mu(A_1 + B_1) = \mu A_1 + \mu B_1 \geqq \mu^* fA + \mu^* fB$, and so $\nu^*(A + B) \geqq \nu^* A + \nu^* B$. The inverse inequality follows from M2, and so M4 is satisfied. To verify condition M5, we have $\nu^* A = \mu^* fA = \inf_{G \supset fA} \mu G$. But whenever G is an open set containing fA, $f^{-1}G$ is an open set in R containing A. Hence $\nu^* A \geqq \inf_{G \supset A} \nu G$. Again the inverse inequality follows from M2, and the verification that ν is a LS measure is complete. Furthermore, ν is positive for non-void open sets in R, because the image of such a set contains a neighborhood of a regular point of E. Finally, it is evident that single points have measure zero, and by hypothesis $\nu BdR = \mu fBdR = 0$. Hence ν is an r-dimensional LS measure in R.

LEMMA 3: *Let μ be an r-dimensional LS measure in a regularly connected polyhedron E, $r \geqq 1$. It is possible to represent E as the continuous image of a rectangular r-cell R under a map f which is a homeomorphism up to the boundary, and which is such that $\mu^* fA = m^* A$ for all $A \subset R$, where m^* denotes ordinary Lebesgue outer measure in R.*

E is homeomorphic to the polyhedron E_1 of a euclidean complex K_1. It is possible to choose the correspondence in such a way that all $(r - 1)$-dimensional simplexes of K_1 correspond to sets of measure zero. Because if μ_1 is the measure in E_1 corresponding to μ under any homeomorphism, the singular $(r - 1)$-simplexes have measure zero by hypothesis, and a suitably chosen automorphism of E_1 will displace all regular $(r - 1)$-simplexes into sets of μ_1-measure zero. (First displace the interiors of all 1-simplexes not contained in the set of singular points to nearby positions having measure zero. Then do the same for 2-simplexes, and so on until all regular $(r - 1)$-simplexes have been displaced to sets having μ_1-measure zero. The automorphism need displace only regular points of E_1 and may leave all vertices of K_1 fixed.[28]) We may therefore suppose that E is the polyhedron of a euclidean complex K and that all lower dimensional simplexes of K have μ-measure zero. By Lemma 1 there exists a continuous map of a convex cell Z onto E which is a homeomorphism up to the boundary and is a simplicial map of a certain subdivision of Z onto K. Let us combine this map with a homeomorphism of Z onto a rectangular r-cell R, of volume $mR = \mu E$,

[28] More generally, any set of first category can be displaced to a set of measure zero by an automorphism. See our joint paper, *On the equivalence of any set of first category to a set of measure zero*. Fund. Math. **31** (1938) 201–206.

and let f denote the resulting map of R onto E. Then f is a homeomorphism up to the boundary, and since BdR corresponds to certain $(r-1)$-dimensional simplexes of K, we see that $\mu fBdR = 0$. By Lemma 2, the function $\nu^*A = \mu^*fA$ defines an r-dimensional LS measure in R, and also we have $\nu R = mR$. In Part II (Theorem 2) it will be shown that for any such measure there exists an automorphism h of R such that $\nu^*A = m^*hA$ for all $A \subset R$. Assuming this result for the moment, we conclude that fh^{-1} is a continuous map of R onto E which is a homeomorphism up to the boundary, and that $m^*A = \mu^*fh^{-1}A$ for all $A \subset R$, as required.

LEMMA 4: *Let f be a map of R onto E with the properties stated in Lemma 3. Any Lebesgue-measure-preserving automorphism T of R that leaves the boundary fixed corresponds to a μ-measure-preserving automorphism fTf^{-1} of E. On the other hand, any μ-measure-preserving automorphism T of E defines a transformation $f^{-1}Tf$ in R which is a Lebesgue-measure-preserving homeomorphism of the open set $R_0 \cdot f^{-1}T^{-1}fR_0$ onto the open set $R_0 \cdot f^{-1}TfR_0$, both of which are contained in the interior R_0 of R and have measure equal to mR.*

To prove the first assertion we first show that fTf^{-1} is a 1:1 map of E onto itself. Consider any point p in fR_0. Then $f^{-1}p$ is single-valued, and therefore also $fTf^{-1}p$. If $p \notin fR_0$, then $p \notin fBdR$, and so $fTf^{-1}p = p$. Any point of $fBdR$ is its own image, and any point $p \notin fR_0$ is the image of $fT^{-1}f^{-1}p$. Hence fTf^{-1} is a 1:1 map of E onto itself, with inverse $fT^{-1}f^{-1}$. Both these transformations carry closed sets into closed sets, and so they are automorphisms. That fTf^{-1} preserves μ-measure is immediate, since $\mu^*A = m^*f^{-1}A = m^*Tf^{-1}A = \mu^*fTf^{-1}A$ for every set $A \subset E$.

To prove the second assertion of the lemma, consider any $p \in R_0 \cdot f^{-1}T^{-1}fR_0$. Then $Tfp \in fR_0$ and so $f^{-1}Tfp$ is single-valued on this domain. Its inverse is $f^{-1}T^{-1}f$ with domain $R_0 \cdot f^{-1}TfR_0$. Both of these are continuous throughout these domains. Thus $f^{-1}Tf$ maps the open set $R_0 \cdot f^{-1}T^{-1}fR_0$ homeomorphically onto $R_0 \cdot f^{-1}TfR_0$, and does so in a measure-preserving manner, as follows from the measure-preserving properties of f and T.

The partial correspondence between measure-preserving automorphisms of E and R described in Lemma 4 makes it possible to reduce the proof of Theorem 1 to the following lemma concerning transformations in R and ordinary Lebesgue measure.

LEMMA 5: *Let R_0 be the interior of a rectangular r-cell R, $r \geqq 2$, and let T be a measure-preserving homeomorphism of an open set $G \subset R_0$ onto an open set $TG \subset R_0$, where $mG = mR$, and let $\sigma_1, \cdots, \sigma_N$ be the cells of any given dyadic subdivision. There exist arbitrarily small automorphisms h_1 and h_2 of R that leave the boundary fixed such that h_1Th_2 is a measure-preserving homeomorphism of $h_2^{-1}G$, which has measure mR, onto h_1TG, and such that under this transformation a certain closed set F is transformed in the following manner: There exists a positive integer K such that the first KN images of F under iteration of h_1Th_2 are disjoint and exactly K are contained in the interior of each cell σ_i and contain exactly half the measure of σ_i.*

We shall prove this lemma in §7. We proceed to show that from it we can deduce

Theorem 1, as regards category. Let f be the map of R onto E defined in Lemma 3, and let $\sigma_1, \sigma_2, \cdots$, be an enumeration of all dyadic cells in R. If σ_i and σ_j belong to the same dyadic subdivision, let $E_{i,j}$ be the set of all automorphisms $T \in M[E, \mu]$ such that for some Borel set A we have $TA = A$, $\mu(A \cdot f\sigma_i) > \frac{3}{4}\mu f\sigma_i$, and $\mu(A \cdot f\sigma_j) < \frac{1}{4}\mu f\sigma_j$, otherwise undefined. Every metrically intransitive automorphism in $M[E, \mu]$ belongs to one of the sets $E_{i,j}$. Because if T is metrically intransitive, there exists a measurable set, and therefore a Borel set A, such that $TA = A$ and $0 < \mu A < \mu E$. Therefore $0 < mf^{-1}A < mR$. Let p, q be points of R at which the Borel set $f^{-1}A$ has metric density 1 and 0 respectively. It follows from Lebesgue's density theorem[29] that in any sufficiently fine dyadic subdivision the cells σ_i and σ_j that contain p and q respectively are such that $m(f^{-1}A \cdot \sigma_i) > \frac{3}{4}m\sigma_i$ and $m(f^{-1}A \cdot \sigma_j) < \frac{1}{4}m\sigma_j$. Therefore $\mu(A \cdot f\sigma_i) > \frac{3}{4}\mu f\sigma_i$ and $\mu(A \cdot f\sigma_j) < \frac{1}{4}\mu f\sigma_j$, so that T belongs to $E_{i,j}$. Thus, to prove that the metrically transitive automorphisms form a residual set in $M[E, \mu]$ it suffices to show that each of the sets $E_{i,j}$ is nowhere dense, because there are only countably many of them and their union contains all metrically intransitive automorphisms in $M[E, \mu]$.

Consider any $T \in M[E, \mu]$ and any set $E_{i,j}$. We shall show that arbitrarily near T there exists an automorphism S which together with a neighborhood lies outside $E_{i,j}$. According to Lemma 4, the transformation $f^{-1}Tf$ is a measure-preserving homeomorphism of the open set $G = R_0 \cdot f^{-1}T^{-1}fR_0$ onto another open set contained in R_0 and likewise having measure mR. Applying Lemma 5 to this transformation, there exist arbitrarily small automorphisms h_1, h_2 of R which leave the boundary fixed, such that $h_1f^{-1}Tfh_2$ is a measure-preserving homeomorphism of $h_2^{-1}G$ onto another open set with measure mR. Also there is a closed set F whose first KN images under this transformation are disjoint and equally distributed among the cells $\tau_1, \cdots, \tau_N$ of the dyadic subdivision to which σ_i and σ_j belong, and contain half their measure. By Lemma 4, the transformations fh_1f^{-1} and fh_2f^{-1} are automorphisms of E. Hence $S = (fh_1f^{-1})T(fh_2f^{-1})$ is an automorphism of E, and from the properties of $h_1f^{-1}Tfh_2$ it follows that S transforms the closed set fF in such a way that its first KN images are disjoint and equally distributed among $f\tau_1, \cdots, f\tau_N$, and contain half their measure. Now consider any Borel set A invariant under S. Let α be the fraction of the measure of fF contained in A. Suppose $\alpha \leqq \frac{1}{2}$, then the K disjoint images of fF in σ_i contain the same fraction of the measure of A, and therefore $mA\sigma_i \leqq \left(\frac{1}{2} + \frac{\alpha}{2}\right)m\sigma_i \leqq \frac{3}{4}m\sigma_i$, so that S does not belong to $E_{i,j}$. On the other hand, if $\alpha \geqq \frac{1}{2}$ the K disjoint images of fF in σ_j contain measure $\frac{\alpha}{2}m\sigma_j$, and therefore $mA\sigma_j \geqq \frac{\alpha}{2}m\sigma_j \geqq \frac{1}{4}m\sigma_j$, so that in this case also S does not belong to $E_{i,j}$. Furthermore, the first KN images of fF under any automorphism sufficiently close to S will be distributed among $f\tau_1, \cdots, f\tau_N$ in the same

[29] See e.g. Hobson, *The theory of functions of a real variable*, 2nd edition, Cambridge 1921, vol. 1, p. 181.

way as under S, so that the same reasoning shows that a whole neighborhood of S lies outside $E_{i,j}$. But h_1 and h_2 were arbitrarily small, hence also the automorphisms fh_1f^{-1} and fh_2f^{-1}, so that S is arbitrarily close to T. Hence $E_{i,j}$ is nowhere dense in $M[E, \mu]$.

It will be observed that S may be written in the form S_1T, where S_1 is the measure-preserving transformation ST^{-1}. But S_1 may also be written $(fh_1f^{-1})T(fh_2f^{-1})T^{-1}$, from which it follows that S_1 leaves all singular points fixed. We have therefore proved the more precise result that $E_{i,j}$ is nowhere dense with respect to every coset of the closed normal subgroup of automorphisms that leave all singular points fixed. Since these cosets are complete spaces, we may state Theorem 1 in the following slightly stronger form which we shall need later.

COROLLARY: *For every $T \in M[E, \mu]$ there exists an arbitrarily small automorphism $S \in M[E, \mu]$, that leaves all singular points fixed, such that ST is metrically transitive with respect to μ.*

II. MEASURES TOPOLOGICALLY EQUIVALENT TO LEBESGUE MEASURE

3. The Fundamental Theorem

Let μ be any measure function in a space, and h an automorphism of the space. The function μhA, considered as defined for all sets such that hA is measurable, is easily seen to be a measure function. It would naturally be described as a measure automorphic to μ, (or equivalent to μ under automorphism). It is easily seen that any measure automorphic to a LS measure is again a LS measure, and that the same transformation also effects a correspondence between their outer measures. Likewise, any measure automorphic to an r-dimensional measure is r-dimensional. The object of the present Part will be to obtain a simple characterization of measures automorphic to Lebesgue measure. The basic result, stated as Theorem 2, was originally proposed by one of us in 1936, in connection with some other group-theoretic investigations, and a proof was obtained at that time by J. von Neumann, but was not published. The present proof, based on somewhat different considerations, was worked out subsequently. The result is here published for the first time.

THEOREM 2: *In order that a measure μ in a rectangular r-cell R, $r \geqq 1$, be automorphic to Lebesgue measure it is necessary and sufficient that it be an r-dimensional Lebesgue-Stieltjes measure and that $\mu R = mR$. The correspondence can always be effected by an automorphism that leaves the boundary of R fixed.*

An equivalent formulation, from the standpoint of the Carathéodory theory, runs as follows.

THEOREM 2_1: *In order that a function μ^* defined for all subsets of R be automorphic to the Lebesgue outer measure m^* it is necessary and sufficient that it satisfy conditions $M1$ to $M5$, and also*

M6: $\mu^*G > 0$ *if G is non-void, open*

M7: $\mu^*p = 0$ *for every point p*

M8: $\mu^*BdR = 0$

and finally that $\mu^*R = mR$. If μ^* satisfies these conditions there exists an automorphism h of R such that $\mu^*A = m^*hA$ for every $A \subset R$, and such that h leaves the boundary fixed.

The equivalence of the two formulations is evident. Before proving the theorem we shall derive from it a number of corollaries and generalizations. In Part V, §13 to §15 and §17, we shall give some further applications.

COROLLARY 1: *Let E be any regularly connected polyhedron of dimension $r \geqq 1$, and let μ_1 and μ_2 be two r-dimensional LS measures in E such that $\mu_1 E = \mu_2 E$. There exists an automorphism h such that $\mu_1^* A = \mu_2^* hA$ for all $A \subset E$, and such that h leaves all singular points of E fixed. In particular, any r-dimensional LS measure in a regularly connected euclidean polyhedron E is automorphic to the Lebesgue measure in E, provided only that $\mu E = mE$.*

Let ν be the r-dimensional LS measure in E defined by $\nu^*A = \frac{1}{2}(\mu_1^* A + \mu_2^* A)$, and let f be the map of a rectangular r-cell R onto E given by Lemma 3. Then $m^*A = \nu^* fA$, and in particular, $\nu fBdR = 0$, so that we have $\mu_1 fBdR = \mu_2 fBdR = 0$. Let ν_1^* and ν_2^* be the outer measures in R defined by $\nu_1^* A = \mu_1^* fA$ and $\nu_2^* A = \mu_2^* fA$. By Lemma 2, these are both r-dimensional LS measures with $\nu_1 R = \nu_2 R = mR$. Hence, by Theorem 2_1, there exist automorphisms h_1, h_2 of R that leave the boundary fixed, such that $\nu_1^* A = m^* h_1 A$ and $\nu_2^* A = m^* h_2 A$. Then $h = f h_2^{-1} h_1 f^{-1}$ is an automorphism of E that leaves singular points fixed and carries μ_1^* into μ_2^* as required.

COROLLARY 2: *Let E be a regularly connected polyhedron of dimension $r \geqq 1$, and E_1 any homeomorphic euclidean polyhedron. The r-dimensional Lebesgue-Stieltjes measures in E are the same as the measures of the form $C \cdot mhA$, where h is an arbitrary homeomorphism of E onto E_1, C is an arbitrary positive constant, and m denotes Lebesgue measure in E.*

This follows at once from Corollary 1 and the fact that any multiple of a measure homeomorphic to an r-dimensional measure is again an r-dimensional measure.

COROLLARY 3: *Let E_1 and E_2 be homeomorphic regularly connected polyhedra of dimension $r \geqq 1$, and let μ_1 and μ_2 be r-dimensional Lebesgue-Stieltjes measures in E_1 and E_2 respectively, such that $\mu_1 E = \mu_2 E$. Let h be any homeomorphism of E_1 onto E_2. There exists a homeomorphism g of E_1 onto E_2 which carries μ_1 into μ_2 and which is equal to h for all singular points. In particular, there exists a measure-preserving automorphism of a rectangular r-cell which is equal to any given automorphism on the boundary.*

The measure $\nu_1 A = \mu_2 hA$ is an r-dimensional LS measure in E_1, and $\nu_1 E_1 = \mu_1 E_1$. By Corollary 1, there exists an automorphism h_1 of E_1 such that $\mu_1 A = \nu_1 h_1 A$, and such that it leaves singular points fixed. Hence $g = hh_1$ is a homeomorphism of E_1 onto E_2 that takes μ_1 into μ_2 and is equal to h for all singular points.

For later use the following corollary is convenient.

COROLLARY 4: *Let μ be a Lebesgue-Stieltjes measure in a rectangular r-cell R, $r \geqq 2$, $\mu R = mR$, and let L be any straight line segment contained in the interior*

of R such that $\mu L = 0$. Then the automorphism that carries μ into m may be so chosen that it leaves L fixed, as well as the boundary.

To see this, let R_1 and R_2 be rectangular r-cells with R_1 contained in the interior of R_2, and let E be the polyhedron obtained from R_2 by removing the interior of R_1. In the case $r \geqq 2$ it is evident that E is regularly connected. From E we can obtain R by identifying points of the inner boundary that have the same x_1 coordinate, say. That is, we can define a continuous map f of E onto R which carries the inner boundary into L and elsewhere is 1:1. The two measures $\mu_1 A = mfA$ and $\mu_2 A = \mu fA$ are easily verified to be r-dimensional LS measures in E, and $\mu_1 E = \mu_2 E$. By Corollary 1, there exists an automorphism h which leaves the inner and outer boundary of E fixed and carries μ_1 into μ_2. Hence fhf^{-1} is an automorphism of R that carries μ into m and leaves both L and BdR fixed.

The next two corollaries characterize measures automorphic to Lebesgue measure in the whole of euclidean space $E^{(r)}$.

COROLLARY 5: *A function μ^* defined for all subsets of $E^{(r)}$, $r \geqq 2$, is automorphic to Lebesgue outer measure m^* if and only if it satisfies the conditions stated in Theorem 2_1, except that conditions* M1 *and* M8 *are to be replaced by*

M1′: $0 \leqq \mu^* A \leqq \mu^* E^{(r)} = +\infty$

M8′: $\mu^* A < +\infty$ *for every bounded set A.*

The necessity of each of the eight conditions is evident. To prove that they are sufficient, let μ be the measure defined by any such outer measure. Let R_α be the closed cube in $E^{(r)}$ with edges parallel to the axes, center at the origin, and edge length α. Under our hypotheses, the function $f(\alpha) = \mu R_\alpha$ is a finite, strictly increasing function in $0 < \alpha < +\infty$. Such a function can have at most countably many discontinuities, hence we can select a sequence of values $0 < \alpha_1 < \alpha_2 < \cdots$, tending to infinity, at each of which $f(\alpha)$ is continuous. This means that $\mu Bd R_{\alpha_i} = 0$. By hypothesis we have also that $\mu R_{\alpha_i} \to +\infty$ as $i \to \infty$. Let β_i be the edge length such that $mR_{\beta_i} = \mu R_{\alpha_i}$. Then $0 < \beta_1 < \beta_2 < \cdots$, and $\beta_i \to +\infty$ as $i \to \infty$. We can therefore define a radial transformation h_1 of $E^{(r)}$ to carry R_{α_i} into R_{β_i}, $i = 1, 2, \cdots$. Consider the new outer measure $\mu_1^* A = \mu^* h_1^{-1} A$. Then $\mu_1 R_{\beta_i} = mR_{\beta_i}$ and $\mu_1 Bd R_{\beta_i} = 0$, $i = 1, 2, \cdots$. Let $E_1 = R_{\beta_1}$ and for $n > 1$ let E_n be R_{β_n} minus the interior of $R_{\beta_{n-1}}$. Then each E_n is a regularly connected polyhedron, since we have assumed $r \geqq 2$, and μ_1 is an r-dimensional measure in it such that $\mu_1 E_n = mE_n$. By Corollary 1, there exists an automorphism h_2 of E_n such that $\mu_1^* A = m^* h_2 A$, for all $A \subset E_n$. Since h_2 leaves the inner and outer boundaries of E_n fixed, these automorphisms join up to form an automorphism of $E^{(r)}$. In view of the measurability of the sets E_n and the fact that their boundaries have measure zero, we have, for any set $A \subset E^{(r)}$,

$$\mu_1^* A = \sum_{n=1}^{\infty} \mu_1^* A E_n = \sum_{n=1}^{\infty} m^* h_2 (A E_n) = \sum_{n=1}^{\infty} m^* (h_2 A) E_n = m^* h_2 A.$$

Therefore $\mu^* A = m^* h_2 h_1 A$ for all $A \subset E^{(r)}$.

It should be added that in the case $r = 1$ the conditions stated in Corollary 5 are not sufficient, but a complete characterization is obtained by merely adding the requirement that the positive and negative halves of the line both have infinite measure. The function $f(x) = \mu(0 \leqq t \leqq x)$ for $x \geqq 0$, $f(x) = -\mu(x \leqq t \leqq 0)$ for $x < 0$ then defines an automorphism of the line that carries μ into m.

COROLLARY 6: *A measure in $E^{(r)}$, $r \geqq 2$, is automorphic to Lebesgue measure if and only if it is an infinite Lebesgue-Stieltjes measure that is zero for points and positive for non-void open sets.*

Here we may understand by a LS measure one that is derived from a non-negative additive function of an interval. Such a measure is necessarily finite for bounded sets, and so condition M8′ is implicit. The rest then follows from Corollary 5.

4. Proof of Theorem 2

It may be remarked first of all that in the case $r = 1$ Theorem 2 is almost trivial. For suppose R is the unit interval $0 \leqq x \leqq 1$, and let $hx = \mu(0 \leqq t \leqq x)$. Then hx is strictly increasing, in virtue of the additivity of μ and the fact that all subintervals have positive measure. It is also continuous, since points have measure zero. As x describes the interval $(0, 1)$, hx describes the same interval. Hence h is an automorphism of R. By definition, we have $\mu A = mhA$ for every interval A of the form $(0, x)$. From additivity it follows that equality holds for all intervals, and therefore for all open sets. But this implies that $\mu^*A = m^*hA$ for all sets A, by condition M5. Hence h effects the desired transformation, and also leaves the end points fixed.

However, even in the case $r = 2$ the difficulties of the general case already appear. Our proof will be based on a sequence of lemmas whose motivation lies in the idea of securing first that $\mu A = mhA$ for all sets of a division automorphic to a dyadic subdivision. Then h is modified within each of these sets so as to secure equality for the sets of a finer subdivision. Finally, a convergent sequence of such modifications is obtained and the limiting automorphism effects the desired transformation for all sets.

LEMMA 6: *Let μ be any Lebesgue-Stieltjes measure in R that is zero for points and for the boundary, and let α be any number in the interval $0 < \alpha < \mu R$. There exists an open set G contained in the interior of R such that $\mu G = \alpha$.*

Since μR is finite, there can be at most countably many planes parallel to the faces of R that intersect R in sets of positive μ-measure. Hence we can divide R into a finite number of rectangular r-cells $\sigma_1, \cdots, \sigma_N$ of diameter less than $\frac{1}{2}$, whose boundaries all have μ-measure zero. Let i be the least integer such that $\mu(\sigma_1 + \sigma_2 + \cdots + \sigma_i) \geqq \alpha$, and let G_1 be the union of the interiors of $\sigma_1, \sigma_2, \cdots, \sigma_{i-1}$. (Take G_1 equal to the void set in case $i = 1$). Then $\mu G_1 < \alpha$, but $\mu G_1 + \mu\sigma_i \geqq \alpha$. Now consider the cell σ_i, calling it R_1, and divide it into rectangular r-cells $\sigma_1^{(1)}, \sigma_2^{(1)}, \cdots, \sigma_{N_1}^{(1)}$ of diameter less than $\frac{1}{4}$, whose boundaries all have μ-measure zero. Again we find an open set $G_2 \subset R_1$, either void or

consisting of the interiors of some of the cells $\sigma_i^{(1)}$, such that $\mu(G_1 + G_2) < \alpha$, while $\mu(G_1 + G_2) + \mu R_2 \geqq \alpha$, where R_2 is one of the cells $\sigma_i^{(1)}$ and is disjoint to $G_1 + G_2$. Proceeding in this manner, we find disjoint open sets $G_1, G_2, \cdots$ interior to R, and a nested sequence of rectangular r-cells $R_1, R_2, \cdots$ whose diameters tend to zero, such that $\alpha - \mu R_n \leqq \mu(G_1 + G_2 + \cdots + G_n) < \alpha$, $n \geqq 1$. The cells R_n intersect in a point p, and we have $\lim_{n\to\infty} \mu R_n = \mu p = 0$, by hypothesis. Hence $\mu G = \alpha$, where G is the union of the sets G_n.

LEMMA 7: *Let μ be any Lebesgue-Stieltjes measure in R that is zero for points and for the boundary. Let R_1 and R_2 be the two cells obtained by bisecting R perpendicularly to one of its edges. Let α_1 and α_2 be any two positive numbers such that $\alpha_1 + \alpha_2 = \mu R$. There exists an automorphism $h \in H_0[R]$ such that $\mu h R_1 = \alpha_1$ and $\mu h R_2 = \alpha_2$.*

We shall establish the existence of h by a category argument. Let H_1 be the set of automorphisms h in $H_0[R]$ such that $\mu h R_1 \geqq \alpha_1$ and $\mu h R_2 \geqq \alpha_2$. This is a closed set, for if h_n is any sequence of automorphisms in H_1 tending uniformly to h, then any ϵ-neighborhood of hR_i, $i = 1, 2$, contains $h_n R_i$ for all sufficiently large n, and therefore has μ-measure at least α_i. This being true for every $\epsilon > 0$, it follows that $\mu h R_1 \geqq \alpha_1$ and $\mu h R_2 \geqq \alpha_2$. Furthermore, the set H_1 is non-void. For, unless it contains the identity, we have either $\mu R_1 < \alpha_1$ or $\mu R_2 < \alpha_2$. Suppose $\mu R_1 < \alpha_1$. We can deform R by a continuous family of automorphisms $h_\lambda \in H_0[R]$ in such a way that as λ increases, $h_\lambda R_1$ includes more and more of the interior of R_2. At the limiting value λ_0 where $\mu h_\lambda R_1$ first becomes greater than or equal to α_1 we have also $\mu h_\lambda R_2 \geqq \alpha_2$, because $\mu h_{\lambda_0} R_2 = \lim_{\lambda \to \lambda_0^-} \mu h_\lambda R_2$, and for $\lambda < \lambda_0$ we have $\mu h_\lambda R_2 \geqq \mu R - \mu h_\lambda R_1 > \mu R - \alpha_1 = \alpha_2$. Hence the set H_1, as a non-void closed subset of a complete space, is itself a complete space. We shall show that in it the set of automorphisms that fulfill the requirements of the lemma is residual.

Let E_n be the set of all h in H_1 such that $\mu h R_1 \geqq \alpha_1 + \frac{1}{n}$. This set is closed in H_1, since we have already seen that it is closed in $H_0[R]$. To show that it is nowhere dense in H_1 it suffices to show that if $h \in E_n$ there exists an arbitrarily small automorphism $g \in H_0[R]$ such that $hg \in H_1 - E_n$. Suppose that $h \in E_n$ and consider the new measure $\nu A = \mu h A$. Then $\nu R_1 \geqq \alpha_1 + \frac{1}{n}$ and $\nu R_2 \geqq \alpha_2$. Let R_3 denote the $(r-1)$-dimensional face common to R_1 and R_2. Then $\nu R_1 + \nu R_2 - \nu R_3 = \nu R = \alpha_1 + \alpha_2$. Hence $\nu R_1 - \alpha_1 \leqq \nu R_1 - \alpha_1 + \nu R_2 - \alpha_2 = \nu R_3$, and this with the inequality $\nu R_1 \geqq \alpha_1 + \frac{1}{n}$ gives $0 < \nu R_1 - \alpha_1 - \frac{1}{2n} < \nu R_3$. Considered with respect to R_3, ν is a LS measure that is zero for points and for the boundary. Hence, by Lemma 6, there exists a set G, open with respect to R_3 and contained in its interior, such that $\nu G = \nu R_1 - \alpha_1 - \frac{1}{2n}$. Now let g be an automorphism of R that leaves the boundary fixed and also the points

of $R_3 - G$, but let it displace all points of G slightly into the interior of R_1. Such an automorphism can be found arbitrarily close to the identity. Since $gR_2 \supset R_2$, we have $\nu gR_2 \geqq \alpha_2$; and νgR_1 differs only slightly from $\nu R_1 - \nu G$, which is equal to $\alpha_1 + \frac{1}{2n}$. If g is sufficiently near the identity we shall have $\alpha_1 < \nu gR_1 < \alpha_1 + \frac{1}{n}$. Going back to the measure μ, we have $\alpha_1 < \mu hgR_1 < \alpha_1 + \frac{1}{n}$ and $\mu hgR_2 \geqq \alpha_2$. Hence $hg \in H_1 - E_n$, and so E_n is nowhere dense in H_1. Therefore the set $H_1 - \sum_1^\infty E_n$ is residual, that is, the equation $\mu hR_1 = \alpha_1$ holds for a residual set of automorphisms in H_1. The rôle of R_2 being similar, we see that the automorphisms h such that $\mu hR_2 = \alpha_2$ is also residual in H_1. The intersection of these two residual sets is the set of automorphisms having the properties required by the lemma, and therefore such automorphisms exist.

It may be remarked that in the proof of this lemma it is possible to confine attention to automorphisms that involve only a single coordinate, namely that in the direction along which R is bisected. Thus, if desired, the automorphism h can be chosen so as to leave all coordinates unchanged except this one.

LEMMA 8: *Let μ be any Lebesgue-Stieltjes measure in R that is zero for points and for the boundary. Let $\sigma_1, \cdots, \sigma_N$ be the cells of any dyadic subdivision of R, and let $\alpha_1, \cdots, \alpha_N$ be associated positive numbers whose sum is equal to μR. There exists an automorphism $h \in H_0[R]$ such that $\mu h\sigma_i = \alpha_i$, $i = 1, \cdots, N$.*

The proof consists in applying Lemma 7 a finite number of times. The cells of any dyadic subdivision are obtained from R by making a finite number of bisections. Let R_1, R_2 be the cells obtained from the first bisection. Let $R_{11}, R_{12}; R_{21}, R_{22}$ be the cells obtained by bisecting R_1 and R_2. After n bisections we get 2^n cells $R_{i_1 \cdots i_n}$, $i_1, \cdots, i_n = 1, 2$, and for $n = K \equiv \log_2 N$ the cells $R_{i_1 \cdots i_K}$ are the cells $\sigma_1, \cdots, \sigma_N$ renamed. Let $\beta_{i_1 \cdots i_K}$ be the number α_i associated with $R_{i_1 \cdots i_K}$. For $1 \leqq n < K$ let $\beta_{i_1 \cdots i_n} = \sum \beta_{i_1 \cdots i_K}$ summed over $i_{n+1}, \cdots, i_K = 1, 2$. For $n = 1$ we have $\beta_1 + \beta_2 = \mu R$, and by Lemma 7 we can find h_1 such that $\mu h_1 R_i = \beta_i$, $i = 1, 2$. Suppose we have defined h_n such that $\mu h_n R_{i_1 \cdots i_n} = \beta_{i_1 \cdots i_n}$, $i_1, \cdots, i_n = 1, 2$. Form the new measure $\mu_n A = \mu h_n A$. Then μ_n satisfies the conditions of Lemma 7 with respect to each of the cells $R_{i_1 \cdots i_n}$, and since $\beta_{i_1 \cdots i_n 1} + \beta_{i_1 \cdots i_n 2} = \mu_n R_{i_1 \cdots i_n}$, we can find an automorphism $g_{i_1 \cdots i_n}$ in $H_0[R_{i_1 \cdots i_n}]$ such that $\mu_n g_{i_1 \cdots i_n} R_{i_1 \cdots i_{n+1}} = \beta_{i_1 \cdots i_{n+1}}$, $i_{n+1} = 1, 2$. Since $g_{i_1 \cdots i_n}$ leaves the boundary of $R_{i_1 \cdots i_n}$ fixed, these automorphisms join up to form an automorphism $g \in H_0[R]$, and we have $\mu h_n g R_{i_1 \cdots i_{n+1}} = \beta_{i_1 \cdots i_{n+1}}$. Setting $h_{n+1} = h_n g$, the induction hypotheses are again satisfied, and for $n = K$ we get an automorphism $h = h_K$ that fulfills the requirements of the lemma.

LEMMA 9: *Let μ be any Lebesgue-Stieltjes measure in R that is zero for points and for the boundary. There exists an automorphism $h \in H_0[R]$ such that for every dyadic cell σ we have $\mu hBd\sigma = 0$.*

We shall show that the automorphisms having the desired property form a

residual set in $H_0[R]$.[30] Let R_1 be the $(r-1)$-dimensional cell in which R is intersected by any one of the planes used in forming dyadic subdivisions. Let E_n be the set of automorphisms $h \in H_0[R]$ such that $\mu h R_1 \geqq \frac{1}{n}$. This is a closed set. If h is any element of E_n, let $\nu A = \mu h A$. Choose $\epsilon > 0$ so that the ϵ-neighborhood of R_1 has ν-measure less than $\nu R_1 + \frac{1}{n}$. Let $g \in H_0[R]$ be an automorphism that displaces the interior of R_1 into a set disjoint to R_1 but contained in its ϵ-neighborhood. Then, since the intersection of R_1 and gR_1 has ν-measure zero, we have $\nu R_1 + \nu g R_1 = \nu(R_1 + gR_1) \leqq \nu(\epsilon\text{-neighborhood of } R_1) < \nu R_1 + \frac{1}{n}$. Hence $\mu h g R_1 < \frac{1}{n}$, and $hg \in H_0 - E_n$. Since hg is arbitrarily near h, the set E_n is nowhere dense, and $H_0 - \sum_1^\infty E_n$ is residual. The intersection of the residual sets corresponding to each of the countably many planes used in forming dyadic subdivisions is therefore also residual, and the automorphisms belonging to this set have the required property.

In the following lemma we require for the first time that the measures under consideration be positive for non-void open sets.

LEMMA 10: *Let μ, ν be two r-dimensional Lebesgue-Stieltjes measures in R such that $\mu R = \nu R$, and let $\epsilon > 0$ be given. There exist automorphisms g, h in $H_0[R]$ such that for each cell σ of a certain dyadic subdivision of R we have $\mu g \sigma = \nu h \sigma$; $\mu g B d \sigma = \nu h B d \sigma = 0$; diam $g\sigma < \epsilon$, diam $h\sigma < \epsilon$, diam $\sigma < \epsilon$.*

By Lemma 9, there exists an automorphism g_1 such that for every dyadic cell σ we have $\mu g_1 B d \sigma = 0$. Let μ_1 be the measure $\mu_1 A = \mu g_1 A$. Let $\sigma_1, \cdots, \sigma_N$ be the cells of a dyadic subdivision such that diam $\sigma_i < \epsilon$ and diam $g_1\sigma_i < \epsilon$, $i = 1, \cdots, N$. The numbers $\mu_1\sigma_i$ are all positive and their sum is equal to νR. Hence, by Lemma 8, there exists an automorphism h_1 such that $\nu h_1 \sigma_i = \mu_1 \sigma_i$, $i = 1, \cdots, N$. Put $\nu_1 A = \nu h_1 A$. Then since $\nu_1 B d \sigma_i = 0$, we can apply Lemma 9 to each cell σ_i and get an automorphism h_2 that transforms each cell σ_i into itself, such that $\nu_1 h_2 B d \sigma = 0$ for every dyadic cell σ. Let σ_{ij}, $i = 1, \cdots, N$, $j = 1, \cdots, M$, $\sigma_{ij} \subset \sigma_i$, be the cells of a finer dyadic subdivision of R, such that diam $h_1 h_2 \sigma_{ij} < \epsilon$. Since the numbers $\nu_1 h_2 \sigma_{ij}$ are all positive, and $\sum_j \nu_1 h_2 \sigma_{ij} = \mu_1 \sigma_i$, we can apply Lemma 8 to each cell σ_i using the measure μ_1 and get an automorphism g_2 that transforms each cell σ_i into itself, such that $\mu_1 g_2 \sigma_{ij} = \nu_1 h_2 \sigma_{ij}$. That is, $\mu g_1 g_2 \sigma_{ij} = \nu h_1 h_2 \sigma_{ij}$. Since the sum of these MN numbers is equal to μR (or νR), we have also $\mu g_1 g_2 B d \sigma_{ij} = \nu h_1 h_2 B d \sigma_{ij} = 0$. Finally, since $h_1 h_2 \sigma_{ij} \subset h_1 \sigma_{ij}$, $g_1 g_2 \sigma_{ij} \subset g_1 \sigma_{ij}$, and $\sigma_{ij} \subset \sigma_i$, we have also diam $h_1 h_2 \sigma_{ij} < \epsilon$, diam $g_1 g_2 \sigma_{ij} < \epsilon$, and diam $\sigma_{ij} < \epsilon$, and so the automorphisms $g = g_1 g_2$, $h = h_1 h_2$ fulfill all the requirements of the lemma.

LEMMA 11: *Any two r-dimensional Lebesgue-Stieltjes measures μ, ν in R such*

[30] Cf. footnote 28.

that $\mu R = \nu R$ are automorphic to each other under an automorphism that leaves the boundary fixed.

The proof consists in finding a sequence of partitions of R into cells[31] $\sigma_1^{(n)}, \cdots, \sigma_{N_n}^{(n)}$ and two sequences of automorphisms g_n, $h_n \in H_0[R]$, with the following properties, where $\epsilon_n = \frac{1}{2^n} \operatorname{diam} R$.

1° $\operatorname{diam} \sigma_i^{(n)} \leqq \epsilon_n$, $\operatorname{diam} g_n\sigma_i^{(n)} \leqq \epsilon_n$, $\operatorname{diam} h_n\sigma_i^{(n)} \leqq \epsilon_n$

2° $g_n\sigma_i^{(n-1)} = g_{n-1}\sigma_i^{(n-1)}$, $h_n\sigma_i^{(n-1)} = h_{n-1}\sigma_i^{(n-1)}$, $i = 1, \cdots, N_{n-1}$

3° $\mu g_n Bd\sigma_i^{(n)} = \nu h_n Bd\sigma_i^{(n)} = 0$, $i = 1, \cdots, N_n$

4° $\mu g_n\sigma_i^{(n)} = \nu h_n\sigma_i^{(n)}$, $i = 1, \cdots, N_n$

These conditions are satisfied for $n = 0$ if we take g_0, h_0 equal to the identity, and $\sigma^{(0)} = R$. Suppose g_n, h_n; $\sigma_1^{(n)}, \cdots, \sigma_{N_n}^{(n)}$ have been defined so that conditions 1° to 4° are satisfied. Let $\mu_n A = \mu g_n A$ and $\nu_n A = \nu h_n A$. Then for each cell $\sigma_i^{(n)}$ we have $\mu_n\sigma_i^{(n)} = \nu_n\sigma_i^{(n)}$ and $\mu_n Bd\sigma_i^{(n)} = \nu_n Bd\sigma_i^{(n)} = 0$. Hence μ_n and ν_n satisfy the conditions of Lemma 10 relative to $\sigma_i^{(n)}$, and there exist automorphisms g', $h' \in H_0[\sigma_i^{(n)}]$ such that for each cell $\sigma_{ij}^{(n)}$ of a certain dyadic subdivision of $\sigma_i^{(n)}$ we have $\mu_n g'\sigma_{ij}^{(n)} = \nu_n h'\sigma_{ij}^{(n)}$; $\mu_n g' Bd\sigma_{ij}^{(n)} = \nu_n h' Bd\sigma_{ij}^{(n)} = 0$; and $\operatorname{diam} g'\sigma_{ij}^{(n)} < \epsilon$, $\operatorname{diam} h'\sigma_{ij}^{(n)} < \epsilon$, $\operatorname{diam} \sigma_{ij}^{(n)} < \epsilon$, where we suppose ϵ chosen less than ϵ_{n+1}, and so small that both g_n and h_n take sets of diameter less than ϵ into sets of diameter less than ϵ_{n+1}. Then we have $\mu g_n g'\sigma_{ij}^{(n)} = \nu h_n h'\sigma_{ij}^{(n)}$; $\mu g_n g' Bd\sigma_{ij}^{(n)} = \nu h_n h' Bd\sigma_{ij}^{(n)} = 0$; and $\operatorname{diam} g_n g'\sigma_{ij}^{(n)} < \epsilon_{n+1}$, $\operatorname{diam} h_n h'\sigma_{ij}^{(n)} < \epsilon_{n+1}$, $\operatorname{diam} \sigma_{ij}^{(n)} < \epsilon_{n+1}$. The automorphisms g', h' thus defined in each of the cells $\sigma_i^{(n)}$ join up to form automorphisms of R, which we shall also denote by g', h'. Putting $g_{n+1} = g_n g'$, $h_{n+1} = h_n h'$ and taking the totality of cells $\sigma_{ij}^{(n)}$ for $\sigma_1^{(n+1)}, \cdots, \sigma_{N_{n+1}}^{(n+1)}$, we see that conditions 1° to 4° are again satisfied, and the inductive definition is complete.

From conditions 1° to 4° it follows that the sequences g_n, h_n converge in $H_0[R]$ to automorphisms g, h. Because, if $x \in \sigma_i^{(n)}$ then $g_{n+1}x$ and $g_n x$ both belong to $g_n\sigma_i^{(n)}$, by 2°. Hence $|g_{n+1}x - g_n x| < \epsilon_n$,[32] by 1°. Similarly, if $x \in g_{n+1}\sigma_i^{(n)}$ then $g_n^{-1}x$ and $g_{n+1}^{-1}x$ both belong to $\sigma_i^{(n)}$. Hence $|g_{n+1}^{-1}x - g_n^{-1}x| < \epsilon_n$. Therefore $\rho(g_{n+1}, g_n) < \epsilon_n$, and because of our choice of ϵ_n, $\rho(g_{n+k}, g_n) < \epsilon_{n-1}$. Hence the sequence g_n converges in $H_0[R]$, and the sequence h_n for similar reasons. From 2° it follows that $g\sigma_i^{(n)} = g_n\sigma_i^{(n)}$ and $h\sigma_i^{(n)} = h_n\sigma_i^{(n)}$, and therefore that $gBd\sigma_i^{(n)} = g_n Bd\sigma_i^{(n)}$ and $hBd\sigma_i^{(n)} = h_n Bd\sigma_i^{(n)}$. From 3° and 4° we then find that $\mu g\sigma_i^{(n)} = \nu h\sigma_i^{(n)}$ and $\mu gBd\sigma_i^{(n)} = \nu hBd\sigma_i^{(n)} = 0$. Hence the measures $\mu' A = \mu gA$ and $\nu' A = \nu hA$ are equal for all cells $\sigma_i^{(n)}$ and are zero for their boundaries. Now from 1° it follows that any open set G can be represented as the union of a sequence of non-overlapping cells $\sigma_i^{(n)}$, say $G = \sum_{s=1}^{\infty} \sigma_s$. Then, since the intersection of any two cells σ_s has measure zero with respect to μ'

[31] The cells will all be dyadic cells, but not necessarily of the same subdivision.

[32] We use the vector notation $|x - y|$ to denote the euclidean distance between points x and y.

and ν', we have $\mu'G = \sum_{s=1}^{\infty} \mu'\sigma_s = \sum_{s=1}^{\infty} \nu'\sigma_s = \nu'G$. Hence μ'^* and ν'^* agree for all open sets, and therefore for all sets. Thus $\mu^*A = \nu^* h g^{-1} A$ for every set, and the lemma is proved.

Theorem 2 follows at once on taking $\nu = m$.

In the proof of Lemma 7 it was remarked that in that case the automorphism could be chosen so as to involve only a single coordinate. A simple refinement of the proof of Lemma 9 shows that the automorphism there obtained may be taken to be a finite product of automorphisms each of which involves only a single coordinate. The final automorphism in Lemma 11 is obtained by composing such automorphisms and then passing to the limit. Thus it is possible to assert that the correspondence of the two measure functions can be effected by an automorphism that can be approximated uniformly by finite products of automorphisms each of which involves only a single coordinate. This remark is of interest in view of the fact that it is still an open question whether an arbitrary automorphism of R can be so approximated.[33] If true, this would furnish a strong inductive method for proving topological theorems.

Since the transformation can always be effected by an automorphism that is special at least to the extent of leaving the boundary fixed, and perhaps in other respects, as indicated in the last paragraph, it is natural to inquire whether it would suffice to consider only differentiable automorphisms, or ones whose modulus of continuity is otherwise restricted. That this is not the case may be seen as follows.

Consider the r-dimensional unit cube and let h be a radial contraction that leaves the boundary fixed and takes each concentric cube with edge length $2d < \frac{1}{2}$ into the one whose circumscribed sphere has radius $\delta = e^{-\frac{1}{d}}$, and let $\mu A = mh^{-1}A$. Then the μ-measure of any sphere about the center with radius $\delta < e^{-4}$ is greater than the Lebesgue measure of a sphere with radius $d = \left(\log \frac{1}{\delta}\right)^{-1}$. Hence any automorphism that takes μ into m must have a modulus of continuity at the center greater than $\left(\log \frac{1}{\delta}\right)^{-1}$. But this function dominates any of the form $C\delta^{\alpha}$, $\alpha > 0$, and so the automorphism cannot be differentiable at the center, nor can it satisfy even a Lipschitz or Hölder condition there.

III. Proof of Theorem 1

In Part I, the proof of Theorem 1, as regards category, was reduced to Lemma 5, except for an assumption in the proof of Lemma 3 which has now been removed. Lemma 5 is an approximation theorem. To prove it, it will be necessary to devise methods for modifying a given transformation so as to secure control over the distribution of the images of a certain closed set. We shall do

[33] Problem proposed by one of us in Fund. Math. **24** (1935) 324.

this in three main steps. In §5 it is shown how a transformation can be modified so that it will have a periodic point with images equally distributed among the cells of a given dyadic subdivision. In §6 a lemma is derived which, in effect, enables one to obtain a periodic Cantor set, starting with only a periodic point. These two constructions are combined in §7, and the final step consists in an equivalence transformation which expands the Cantor set to one having the required amount of measure. The result is that an arbitrary measure-preserving automorphism of a rectangular r-cell (or, more generally, a transformation in the r-cell that corresponds to a measure-preserving automorphism of another polyhedron as described in Lemma 4) can be modified so that the resulting transformation is measure-preserving and has a periodic Cantor set whose images carry a prescribed fraction of the total measure and are equally distributed among the cells of a given dyadic subdivision. All modifications are effected by left and right multiplication with automorphisms that leave the boundary fixed. Thus Lemma 5 is established, and therefore Theorem 1. In §18 it will be shown that for automorphisms in $M_0[R]$ one can obtain still more precise approximation theorems.

5. Lemmas Concerning Distribution of Finite Sets

LEMMA 12: *Let p, q be any two interior points of a rectangular r-cell R, $r \geqq 2$. There exists an automorphism $T \in M_0[R]$ such that $Tp = q$.*

Observe first that there is no difficulty in defining an automorphism g that carries p into q and leaves BdR fixed. Most simply, join p to the vertices of R and map the resulting cell complex affinely onto the similar complex obtained from q. Thus the only point to the lemma is to show that the transformation can be effected by a measure-preserving automorphism. It is possible to do this directly by a construction involving displacements around closed tubes, but a more elegant proof is based on Theorem 2. Let g be the automorphism just defined, and consider the outer measure $\mu^*A = m^*g^{-1}A$. By Theorem 2_1 there exists an automorphism $h \in H_0[R]$ such that $\mu^*A = m^*hA$. By Corollary 4, h can be so chosen as to leave q fixed, in fact it may leave a line segment through q fixed. The composed automorphism hg then carries p into q, leaves BdR fixed, and since $m^*hgA = \mu^*gA = m^*A$ we see that $hg \in M_0[R]$.

LEMMA 13: *Given two sets of points $p_1, \cdots, p_N$; $q_1, \cdots, q_N$, each consisting of N distinct interior points of a rectangular r-cell R, $r \geqq 2$, and $|p_i - q_i| < \epsilon$, there exists an automorphism $T \in M_0[R]$ such that $Tp_i = q_i$, $i = 1, \cdots, N$, and $\rho(T, I) < \epsilon$.*

In effect, the lemma asserts that any two finite sets of interior points that could possibly be equivalent under automorphism are equivalent under a measure-preserving automorphism whose norm is no larger than it need be. Let L_i denote the straight line segment joining p_i to q_i. Then L_i is contained in the interior of R. Suppose first that these segments are all disjoint. Then we can enclose each in the interior of a rectangular r-cell σ_i so chosen that the cells $\sigma_1, \cdots, \sigma_N$ are disjoint and each has diameter less than ϵ. By the previous

lemma, we can find an automorphism T_i of each cell σ_i which preserves measure and carries p_i into q_i. Since these automorphisms leave the boundary of each cell σ_i fixed, they can be extended to an automorphism T of R by defining T equal to T_i in each cell σ_i and equal to the identity elsewhere. Since T only permutes points within each cell, it moves no point by more than ϵ, and so $\rho(T, I) < \epsilon$. In case any of the segments L_i intersect, we shall join each p_i to q_i by a chain of $k + 1$ points $p_i = p_{i0}, p_{i1}, \cdots, p_{ik} = q_i$ in such a way that $| p_{i,j} - p_{i,j+1} | < \epsilon/k$ and such that for each $1 \leqq j \leqq k$ the segment L_{ij} joining $p_{i,j-1}$ to $p_{i,j}$ are disjoint. This can always be done; in fact, unless two of the segments L_i intersect in a point that divides each in the same ratio, it will suffice to take the points p_{ij} as the points that divide L_i into k equal segments. In the exceptional case a slight displacement of some of the intermediate points will make the segments $L_{1j}, \cdots, L_{Nj}$ disjoint and will not disturb the inequalities $| L_{ij} | < \epsilon/k$. By the argument just given above, we can find an automorphism T_j of norm less than ϵ/k such that $T_j p_{i,j-1} = p_{i,j}$, $i = 1, \cdots, N$. Hence the composed transformation $T = T_k T_{k-1} \cdots T_1$ will carry each p_i into q_i. Furthermore, it will have norm less than ϵ, because the product of k transformations whose norms are less than ϵ/k always has norm less than ϵ, as may be deduced at once from the triangle inequality.

LEMMA 14: *In a rectangular r-cell R, $r \geqq 2$, let D be a set with measure equal to mR. Let U_δ denote the operation of "taking the δ-neighborhood with respect to D," so that $U_\delta A$ means the set of all points in D at distance less than δ from A. To each $\delta > 0$ corresponds a positive integer λ with the following property: If T is any* 1:1 *measure-preserving transformation of D onto itself and p is any point of D then $(TU_\delta)^\lambda p = D$.*

It is to be emphasized that λ depends only upon δ and the dimension number r, is independent of the transformation T and of the set D. The transformation need not be assumed to be an automorphism, although it will be in the only application we shall make.

Essentially, the proof consists in the observation that the composite operation TU_δ operating on any non-void set either increases its measure by at least a certain minimum amount or else yields the whole domain D. The important thing is to obtain a uniform estimate of this minimum amount.

The integer λ is determined as follows. Let η be the minimum volume of the intersection of the δ-neighborhoods of any two points of R at distance δ apart. This number η is not simply the common volume of two such spheres, because when the two points are near the boundary of R part of the intersection of their δ-spheres may fall outside R. However, we are concerned only with the fact that $\eta > 0$, and a lower bound to η is furnished by the number $V/2^r$, where V is the r-dimensional volume of a sphere of radius $\delta/2$. This is because the intersection of the δ-neighborhoods of any two points at distance δ certainly contains the $\delta/2$-neighborhood of the point of R midway between them, and the minimum volume of the $\delta/2$-neighborhood of any point of R is $V/2^r$, that being the volume of the $\delta/2$-neighborhood of a vertex of R. We shall take for λ the least integer greater than mR/η.

Now consider the composite operation TU_δ. We shall show that λ iterations of this operation, starting from any point of D, yields the whole of D. For let us suppose that $(TU_\delta)^\lambda p$ is not equal to D, and consider any integer n, $0 \leqq n < \lambda$. The set $(TU_\delta)^n p$ is then not δ-dense in D, and so there is a point q in R whose distance from it is exactly δ. Let q' be a limit point of the set $(TU_\delta)^n p$ at distance δ from q. The operation U_δ applied to the set $(TU_\delta)^n p$ therefore adjoins at least the common part in D of the δ-neighborhoods of q and q', and therefore increases the measure of the set by at least η. That is, $mU_\delta(TU_\delta)^n p \geqq m(TU_\delta)^n p + \eta$. Hence, recalling that T preserves measure, $m(TU_\delta)^{n+1} p \geqq m(TU_\delta)^n p + \eta$. This being true for $n = 0, 1, \cdots, \lambda - 1$, it follows that $m(TU_\delta)^\lambda p \geqq \lambda\eta$, which is impossible, since by definition $\lambda\eta > mR$.

LEMMA 15: *Let T be any measure-preserving automorphism of a set D contained in the interior of an r-cell R, $r \geqq 2$, with $mD = mR$. There exist arbitrarily small measure-preserving automorphisms S_1, $S_2 \in M_0[R]$ such that under the transformation $S_2 S_1 T S_2^{-1}$ the centers of the cells of a certain arbitrarily fine dyadic subdivision of R undergo a cyclic permutation.*

This lemma is in a sense the key to the present problem. It enables one to gain control over the successive images of a certain point, in fact, it secures their equal distribution among the cells of a certain dyadic subdivision. Since the proof is somewhat involved, it may be well to sketch the idea in advance. It is first shown that one can find points $p_1, \cdots, p_K$ whose first L images under T constitute a set which is "nearly" uniformly distributed among the cells of a given dyadic subdivision. This much is deduced from the ergodic theorem. (This is the only place in the entire proof of Theorem 1 where the ergodic theorem of Birkhoff is used.) But some of the points of this set may coincide. The next part of the proof consists in modifying T by composition with a small transformation S_1 to secure that a similarly distributed set of points are all distinct and *also* are linked together to form a single cycle. In joining the chains together, it is necessary to add more points, but it is secured that only a "few" points are added. The result is that under $S_1 T$ there is a cycle consisting of points $q_1, \cdots, q_s$ of which KL are distributed exactly like the first L images of $p_1, \cdots, p_K$. It is further secured that the number s is equal to the number of cells in a finer subdivision, and that the points $q_1, \cdots, q_s$ are sufficiently nearly uniformly distributed that they can be set in correspondence with the centers of the cells of this finer subdivision in such a way that corresponding points are close together. By a final transformation S_2 the points $q_1, \cdots, q_s$ are carried into these centers, so that the latter constitute a cycle under $S_2 S_1 T S_2^{-1}$.

Proceeding now to a more precise formulation of this proof, let δ be an arbitrary positive number, and let $\sigma_1, \cdots, \sigma_N$ be the cells of a dyadic subdivision with diameter less than δ, say $N = 2^{\alpha r}$. Let them be numbered in such a way that σ_i has at least one vertex in common with σ_{i+1}, $1 \leqq i < N$. This is easily seen to be possible. Let $f_i(p)$ be the characteristic function of the interior of σ_i. The ergodic theorem asserts that the limit $f_i^*(p) = \lim_{n \to \infty} \frac{1}{n} \sum_{\nu=1}^{n} f_i(T^\nu p)$

lemma, we can find an automorphism T_i of each cell σ_i which preserves measure and carries p_i into q_i. Since these automorphisms leave the boundary of each cell σ_i fixed, they can be extended to an automorphism T of R by defining T equal to T_i in each cell σ_i and equal to the identity elsewhere. Since T only permutes points within each cell, it moves no point by more than ϵ, and so $\rho(T, I) < \epsilon$. In case any of the segments L_i intersect, we shall join each p_i to q_i by a chain of $k + 1$ points $p_i = p_{i0}, p_{i1}, \cdots, p_{ik} = q_i$ in such a way that $| p_{i,j} - p_{i,j+1} | < \epsilon/k$ and such that for each $1 \leqq j \leqq k$ the segment L_{ij} joining $p_{i,j-1}$ to $p_{i,j}$ are disjoint. This can always be done; in fact, unless two of the segments L_i intersect in a point that divides each in the same ratio, it will suffice to take the points p_{ij} as the points that divide L_i into k equal segments. In the exceptional case a slight displacement of some of the intermediate points will make the segments $L_{1j}, \cdots, L_{Nj}$ disjoint and will not disturb the inequalities $| L_{ij} | < \epsilon/k$. By the argument just given above, we can find an automorphism T_j of norm less than ϵ/k such that $T_j p_{i,j-1} = p_{i,j}$, $i = 1, \cdots, N$. Hence the composed transformation $T = T_k T_{k-1} \cdots T_1$ will carry each p_i into q_i. Furthermore, it will have norm less than ϵ, because the product of k transformations whose norms are less than ϵ/k always has norm less than ϵ, as may be deduced at once from the triangle inequality.

LEMMA 14: *In a rectangular r-cell R, $r \geqq 2$, let D be a set with measure equal to mR. Let U_δ denote the operation of "taking the δ-neighborhood with respect to D," so that $U_\delta A$ means the set of all points in D at distance less than δ from A. To each $\delta > 0$ corresponds a positive integer λ with the following property: If T is any* 1:1 *measure-preserving transformation of D onto itself and p is any point of D then $(TU_\delta)^\lambda p = D$.*

It is to be emphasized that λ depends only upon δ and the dimension number r, is independent of the transformation T and of the set D. The transformation need not be assumed to be an automorphism, although it will be in the only application we shall make.

Essentially, the proof consists in the observation that the composite operation TU_δ operating on any non-void set either increases its measure by at least a certain minimum amount or else yields the whole domain D. The important thing is to obtain a uniform estimate of this minimum amount.

The integer λ is determined as follows. Let η be the minimum volume of the intersection of the δ-neighborhoods of any two points of R at distance δ apart. This number η is not simply the common volume of two such spheres, because when the two points are near the boundary of R part of the intersection of their δ-spheres may fall outside R. However, we are concerned only with the fact that $\eta > 0$, and a lower bound to η is furnished by the number $V/2^r$, where V is the r-dimensional volume of a sphere of radius $\delta/2$. This is because the intersection of the δ-neighborhoods of any two points at distance δ certainly contains the $\delta/2$-neighborhood of the point of R midway between them, and the minimum volume of the $\delta/2$-neighborhood of any point of R is $V/2^r$, that being the volume of the $\delta/2$-neighborhood of a vertex of R. We shall take for λ the least integer greater than mR/η.

Now consider the composite operation TU_δ. We shall show that λ iterations of this operation, starting from any point of D, yields the whole of D. For let us suppose that $(TU_\delta)^\lambda p$ is not equal to D, and consider any integer n, $0 \leqq n < \lambda$. The set $(TU_\delta)^n p$ is then not δ-dense in D, and so there is a point q in R whose distance from it is exactly δ. Let q' be a limit point of the set $(TU_\delta)^n p$ at distance δ from q. The operation U_δ applied to the set $(TU_\delta)^n p$ therefore adjoins at least the common part in D of the δ-neighborhoods of q and q', and therefore increases the measure of the set by at least η. That is, $mU_\delta(TU_\delta)^n p \geqq m(TU_\delta)^n p + \eta$. Hence, recalling that T preserves measure, $m(TU_\delta)^{n+1} p \geqq m(TU_\delta)^n p + \eta$. This being true for $n = 0, 1, \cdots, \lambda - 1$, it follows that $m(TU_\delta)^\lambda p \geqq \lambda\eta$, which is impossible, since by definition $\lambda\eta > mR$.

LEMMA 15: *Let T be any measure-preserving automorphism of a set D contained in the interior of an r-cell R, $r \geqq 2$, with $mD = mR$. There exist arbitrarily small measure-preserving automorphisms S_1, $S_2 \in M_0[R]$ such that under the transformation $S_2 S_1 T S_2^{-1}$ the centers of the cells of a certain arbitrarily fine dyadic subdivision of R undergo a cyclic permutation.*

This lemma is in a sense the key to the present problem. It enables one to gain control over the successive images of a certain point, in fact, it secures their equal distribution among the cells of a certain dyadic subdivision. Since the proof is somewhat involved, it may be well to sketch the idea in advance. It is first shown that one can find points $p_1, \cdots, p_K$ whose first L images under T constitute a set which is "nearly" uniformly distributed among the cells of a given dyadic subdivision. This much is deduced from the ergodic theorem. (This is the only place in the entire proof of Theorem 1 where the ergodic theorem of Birkhoff is used.) But some of the points of this set may coincide. The next part of the proof consists in modifying T by composition with a small transformation S_1 to secure that a similarly distributed set of points are all distinct and *also* are linked together to form a single cycle. In joining the chains together, it is necessary to add more points, but it is secured that only a "few" points are added. The result is that under $S_1 T$ there is a cycle consisting of points $q_1, \cdots, q_s$ of which KL are distributed exactly like the first L images of $p_1, \cdots, p_K$. It is further secured that the number s is equal to the number of cells in a finer subdivision, and that the points $q_1, \cdots, q_s$ are sufficiently nearly uniformly distributed that they can be set in correspondence with the centers of the cells of this finer subdivision in such a way that corresponding points are close together. By a final transformation S_2 the points $q_1, \cdots, q_s$ are carried into these centers, so that the latter constitute a cycle under $S_2 S_1 T S_2^{-1}$.

Proceeding now to a more precise formulation of this proof, let δ be an arbitrary positive number, and let $\sigma_1, \cdots, \sigma_N$ be the cells of a dyadic subdivision with diameter less than δ, say $N = 2^{\alpha r}$. Let them be numbered in such a way that σ_i has at least one vertex in common with σ_{i+1}, $1 \leqq i < N$. This is easily seen to be possible. Let $f_i(p)$ be the characteristic function of the interior of σ_i. The ergodic theorem asserts that the limit $f_i^*(p) = \lim_{n\to\infty} \frac{1}{n} \sum_{\nu=1}^{n} f_i(T^\nu p)$

exists for almost all p, and that $\int_D f_i^*(p)\,dm(p) = \int_D f_i(p)\,dm(p) = 1/N$, provided we suppose that $mR = 1$. We shall show first that we can select a finite set of points $p_1, \cdots, p_K$ such that the average of each of the functions $f_i^*(p)$ over this set is nearly equal to its integral, that is, $\left|\frac{1}{K}\sum_{k=1}^{K} f_i^*(p_k) - \frac{1}{N}\right| < \eta$. Here η may be any positive number, but for our purpose it will suffice to take $\eta = 1/N^2$. Let J be an integer greater than $2/\eta$, and for each set of positive integers $i_1, \cdots, i_N$ not exceeding J let $A_{i_1\cdots i_N}$ denote the set of points p in D such that $\frac{i_j - 1}{J} < f_j^*(p) \leqq \frac{i_j}{J}$ for all integers $1 \leqq j \leqq N$. Let $p_{i_1\cdots i_N}$ be any point of $A_{i_1\cdots i_N}$, provided this set has positive measure, otherwise undefined. Then, since on $A_{i_1\cdots i_N}$ the functions $f_j^*(p)$ differ from $f_j^*(p_{i_1\cdots i_N})$ by not more than $1/J$, we have $|1/N - \sum' f_j^*(p_{i_1\cdots i_N})\cdot mA_{i_1\cdots i_N}| \leqq 1/J < \eta/2$, where $\sum'$ denotes summation over all sets of indices $i_1, \cdots, i_N$ such that $mA_{i_1\cdots i_N} > 0$. Now let $r_{i_1\cdots i_N}$ be rational numbers such that $|1/N - \sum' f_j^*(p_{i_1\cdots i_N})r_{i_1\cdots i_N}| < \eta/2$ and $\sum' r_{i_1\cdots i_N} = 1$. Such numbers exist, since $\sum' mA_{i_1\cdots i_N} = mR = 1$. Let K be a common denominator of all the numbers $r_{i_1\cdots i_N}$, and select $Kr_{i_1\cdots i_N}$ points from each set $A_{i_1\cdots i_N}$ that has positive measure. Call these points $p_1, \cdots, p_K$. Then we have

$$\left|\sum' f_j^*(p_{i_1\cdots i_N})r_{i_1\cdots i_N} - \frac{1}{K}\sum_{k=1}^{K} f_j^*(p_k)\right| \leqq \frac{1}{J} < \eta/2$$

and therefore $|1/N - 1/K\sum_{k=1}^{K} f_j^*(p_k)| < \eta$, $j = 1, \cdots, N$. We may suppose the points $p_1, \cdots, p_K$ so chosen that no image falls on the boundary of a cell σ_j, since this means avoiding only a set of measure zero in making the selection. Recalling that $f_j^*(p_k) = \lim_{n\to\infty} \frac{1}{n}\sum_{\nu=1}^{n} f_j(T^\nu p_k)$, we see that for all sufficiently large n we have $\left|1/N - 1/K\sum_{k=1}^{K}\frac{1}{n}\sum_{\nu=1}^{n} f_j(T^\nu p_k)\right| < \eta$, $j = 1, \cdots, N$. Hence for all sufficiently large n we have also

$$\left|\frac{1}{N} - \frac{1}{Kn}\sum_{k=1}^{K}\sum_{\nu=1}^{n} f_j(T^\nu p_k)\right| + \frac{2(\lambda+1)}{n} < \eta \tag{1}$$

where λ is the integer corresponding to δ determined in Lemma 14. Let L be such a value of n, and in addition let it be so chosen that for some integer $\beta > \alpha$ we have $2^{\beta r} - K \leqq K(L + \lambda) < 2^{\beta r}$. Such a number L can be found since the numbers $K(n + \lambda)$ include all multiples of K greater than some number. If we let $s = 2^{\beta r}$ and $K_1 = s - K(L + \lambda)$, then $0 \leqq K_1 \leqq K$.

We now proceed to select s distinct points $q_1, \cdots, q_s$ from D in such a way that Tq_i is within distance δ of q_{i+1}, $1 \leqq i \leqq s$, (letting $q_{s+1} = q_1$), and such that KL of these points are interior to the same cells σ_j as corresponding points

$T^\nu p_k$, $1 \leqq \nu \leqq L$, $1 \leqq k \leqq K$. To secure this, we proceed step by step as follows. Take $q_1 = Tp_1$. For $1 < i \leqq L$, take q_i in the same cell as $T^i p_1$ and such that Tq_i is in the same cell as $T^{i+1}p_1$. By Lemma 14 applied to T^{-1}, $q_L \in (T^{-1}U_\delta)^\lambda p_2$. Hence $Tq_L \in U_\delta(T^{-1}U_\delta)^{\lambda-1}p_2$. Choose q_{L+1} in $(T^{-1}U_\delta)^{\lambda-1}p_2$ within distance δ of Tq_L. For $1 < i \leqq \lambda - 1$, take $q_{L+i} \in (T^{-1}U_\delta)^{\lambda-i}p_2$ within distance δ of Tq_{L+i-1}. Then $Tq_{L+\lambda-1}$ is within distance δ of p_2. Take $q_{L+\lambda}$ within distance δ of $Tq_{L+\lambda-1}$ and such that $Tq_{L+\lambda}$ is in the same cell as Tp_2. For $1 \leqq i \leqq L$, take $q_{L+\lambda+i}$ in the same cell as $T^i p_2$ and such that $Tq_{L+\lambda+i}$ is in the same cell as $T^{i+1}p_2$. The next λ points are then chosen to lead up to p_3, similarly to the way in which q_{L+i}, $1 \leqq i \leqq \lambda$, were chosen leading up to p_2. Eventually, the point $q_{(K-1)(L+\lambda)+L}$ is chosen in the same cell as $T^L p_K$. We then close the cycle by a chain of $\lambda + K_1$ points (instead of λ points as in the preceding cases) in order to get a cycle with exactly s points. This can be done since $q_{(K-1)(L+\lambda)+L} \in (T^{-1}U_\delta)^{\lambda+K_1}p_1$. At the last step we have $q_{s-1} \in T^{-1}U_\delta p_1$ and choose q_s near enough to p_1 so that it is within distance δ of Tq_{s-1} and so that Tq_s is within distance δ of $q_1 = Tp_1$. At every step the point to be selected can be chosen arbitrarily from an open sphere with respect to D. Hence there is no difficulty in avoiding points previously chosen and thus securing that the points $q_1, \cdots, q_s$ are distinct, and therefore also the points $Tq_1, \cdots, Tq_s$. We see that the KL points $q_1, q_2, \cdots, q_L, q_{L+\lambda+1}, \cdots, q_{L+\lambda+L}, \cdots, q_{(K-1)(L+\lambda)+1}, \cdots, q_{(K-1)(L+\lambda)+L}$ are respectively interior to the same cells σ_j as the points $Tp_1, T^2p_1, \cdots, T^Lp_1, Tp_2, \cdots, T^Lp_2, \cdots, Tp_K, \cdots, T^Lp_K$. Finally, it is seen that q_{i+1} is always within distance δ of Tq_i, since they either lie in the same cell, or were specifically chosen so as to be within this distance.

In virtue of Lemma 13, we can find an automorphism $S_1 \in M_0[R]$, with norm less than δ, which carries Tq_i into q_{i+1}, $1 \leqq i \leqq s$. Under the composed transformation S_1T the points $q_1, \cdots, q_s$ thus form a single cycle. Furthermore, since KL of these lie in the same cells as corresponding points $T^\nu p_k$, it follows that

$$\left|\frac{1}{N} - \frac{1}{s}\sum_{i=1}^{s} f_j(q_i)\right| \leqq \left|\frac{1}{N} - \frac{1}{s}\sum_{k=1}^{K}\sum_{\nu=1}^{L} f_j(T^\nu p_k)\right| + \frac{K\lambda + K_1}{s}$$

$$\leqq \left|\frac{1}{N} - \frac{1}{KL}\sum_{k=1}^{K}\sum_{\nu=1}^{L} f_j(T^\nu p_k)\right| + \frac{s - KL}{s} + \frac{K\lambda + K_1}{s}.$$

From equation (1) it follows that this is less than η, because

$$\frac{s - KL}{s} + \frac{K\lambda + K_1}{s} = \frac{2(K\lambda + K_1)}{s} \leqq \frac{2K(\lambda + 1)}{K(L + \lambda) + K_1} < \frac{2(\lambda + 1)}{L}.$$

Hence the relative number of points $q_1, \cdots, q_s$ in any cell σ_j differs from $1/N$ by less than $\eta = 1/N^2$. Let us now re-number the points $q_1, \cdots, q_s$ counting first all those in σ_1, then those in σ_2, and so on. Call the resulting sequence $q_1', q_2', \cdots, q_s'$. Let $q_1'', q_2'', \cdots, q_s''$ be the centers of the $s = 2^{\beta r}$ cells of the β-th dyadic subdivision, again counting first those in σ_1, then those in σ_2, and

so on. We shall show that corresponding points q_i' and q_i'' always lie either in the same cell σ_j or in adjacent cells. The number of points q_i in any cell σ_j differs from s/N by less than s/N^2. Hence the indices of the points q_i' that lie in σ_j fall between the limits $\frac{(j-1)s}{N} - \frac{js}{N^2}$ and $\frac{js}{N} + \frac{js}{N^2}$, which are contained between $\frac{(j-2)s}{N}$ and $\frac{(j+1)s}{N}$, the limits of the indices of points q_i'' lying in σ_{j-1}, σ_j, and σ_{j+1}. Since these are adjacent cells, the distance from q_i' to q_i'' is less than 2δ, and by Lemma 13 we can find an automorphism $S_2 \in M_0[R]$ with norm less than 2δ which carries q_i' into q_i'', $1 \leqq i \leqq s$. The points q_i'' then constitute a single cycle under the transformation $S_2 S_1 T S_2^{-1}$. Since S_1 and S_2 are both arbitrarily near the identity, the lemma is established.

6. Equivalence of Cantor sets under automorphism of the containing space

LEMMA 16: *Let C be a linear Cantor set*[34] *contained in the interior of an r-cell σ_0, which in turn is contained in an r-cell R, $r \geqq 2$, and let f be a homeomorphism of σ_0 onto a subset of R. There exists an automorphism h of R that leaves the boundary fixed and carries fC back into coincidence with C in such a way that $hfp = p$ for every p in C.*

The lemma implies that any topological Cantor set C' contained in the interior of R is equivalent under automorphism of R to a linear Cantor set C provided some homeomorphism of C onto C' can be extended to an r-cell containing C. It should be remarked that such an extension is not always possible even when one of the sets is linear, as is shown by Antoine's example in the 3-dimensional cube (See footnote 15).

The first step in the proof consists in showing that if σ_1 and σ_2 are disjoint r-cells contained in the interior of σ_0, there exists an automorphism $h \in H_0[R]$ such that $hf\sigma_1 \subset \sigma_1$ and $hf\sigma_2 \subset \sigma_2$. This may be seen as follows. Let p_1, p_2 be the centers of σ_1, σ_2 respectively. Then fp_1, fp_2 are distinct interior points of R, and we can find an automorphism $h_1 \in H_0[R]$ which carries them back into p_1, p_2 respectively. Then $h_1 f$ is a homeomorphism of σ_0 onto a subset of R which leaves p_1, p_2 fixed. By continuity, there exist r-cells σ_1', σ_2' about p_1, p_2 such that $h_1 f\sigma_1' \subset \sigma_1$ and $h_1 f\sigma_2' \subset \sigma_2$. Let $g \in H_0[\sigma_0]$ be such that $g\sigma_1 \subset \sigma_1'$ and $g\sigma_2 \subset \sigma_2'$. Such an automorphism is easily defined by dividing σ_0 into two cells containing σ_1, σ_2 in their interiors and then drawing the interior points of each part toward p_1, p_2 respectively. The homeomorphism $h_1 fg$ then carries σ_1 and σ_2 into subsets of themselves. Now consider the transformation h_2 of R equal to fgf^{-1} in $f\sigma_0$ and equal to the identity in the rest of R. That h_2 is an automorphism of R follows from Brouwer's theorem on invariance of region.[35] According to this theorem, $f\sigma_0$ is a closed region, that is, the closure of a connected open set, and interior and boundary points of σ_0 correspond respectively to

[34] That is, a topological Cantor set contained in a straight line segment.

[35] See e.g. AH p. 396.

interior and boundary points of $f\sigma_0$. Hence fgf^{-1} is an automorphism of $f\sigma_0$ that leaves all boundary points fixed, and can therefore be extended to the rest of R by defining it equal to the identity outside. The automorphism $h = h_1h_2$ belongs to $H_0[R]$ and hf carries σ_1 and σ_2 into subsets of themselves, as required.

Now let us choose a family of r-cells $\sigma_{i_1\cdots i_n}; i_1, \cdots, i_n = 1, 2; n = 1, 2, \cdots;$ contained in the interior of σ_0, so as to fulfill the following conditions:

1° For each n, the cells $\sigma_{i_1\cdots i_n}$ are disjoint.

2° $\sigma_{i_1\cdots i_n}$ is contained in the interior of $\sigma_{i_1\cdots i_{n-1}}$.

3° The diameter of $\sigma_{i_1\cdots i_n}$ tends to zero with $1/n$.

4° $C = \prod_{n=1}^{\infty} \sum_{i_1\cdots i_n} \sigma_{i_1\cdots i_n}$.

Such a representation of any linear Cantor set is immediate from its definition. We have shown that there exists an automorphism $h_1 \in H_0[R]$ such that $h_1f\sigma_1 \subset \sigma_1$ and $h_1f\sigma_2 \subset \sigma_2$. Similarly, let h_2 be an automorphism of R that is equal to the identity outside σ_1 and σ_2 and such that $h_2h_1f\sigma_{i_1i_2} \subset \sigma_{i_1i_2}; i_1, i_2 = 1, 2$. This involves only an application of the previous result to the two cells σ_1, σ_2 separately. At the n-th stage, we find an automorphism $h_n \in H_0[R]$ which is equal to the identity outside the cells $\sigma_{i_1\cdots i_{n-1}}$ and is such that $h_nh_{n-1}\cdots h_1f\sigma_{i_1\cdots i_n} \subset \sigma_{i_1\cdots i_n}$. The successive products $h_nh_{n-1}\cdots h_1$ converge uniformly, in virtue of condition 3°, and therefore have for limit a continuous mapping h of R onto itself. That h is likewise 1:1 may be seen as follows. Consider two distinct points p, q outside fC. They are outside the sets $f\sigma_{i_1\cdots i_n}$ for some n. Hence their images under h are the same as under the finite product $h_n\cdots h_1$, and therefore distinct and outside C. On the other hand, h carries distinct points of fC into distinct points of C. In fact, h is equal to f^{-1} on fC, because $hf\sigma_{i_1\cdots i_n} \subset \sigma_{i_1\cdots i_n}$ and therefore hf leaves every point of C fixed. Thus h is a 1:1 continuous map of R onto itself, therefore an automorphism, and it leaves boundary points fixed since it is equal to h_1 there. Furthermore, we have just seen that $hfp = p$ for all p in C.

Before proceeding with our main investigation, which will be resumed beginning with Lemma 17, we shall digress to consider the bearing of the present lemma on the work of Antoine.[36] Given two Cantor sets C and C' situated in a euclidean space $E^{(r)}$ three possibilities are conceivable: (i) there is a homeomorphism of C onto C' that can be extended to the whole space; (ii) there is a homeomorphism that can be extended to a neighborhood of C but none can be extended to the whole space; (iii) no homeomorphism can be extended even to a neighborhood of C. Antoine showed that in the plane only case (i) can arise, indeed that the homeomorphism can be taken equal to the identity outside any given rectangle to which C and C' are interior. But he showed by examples that in 3-space (and therefore in r-space, $r > 3$) all three possibilities can arise, and that case (iii) can be realized even when one of the sets is linear. In his example illustrating case (ii), however, both sets are skew. We shall show

[36] See footnote 15.

that this is necessarily the case, that is, that if either of the sets is plane or linear case (ii) cannot arise in any space $E^{(r)}$. More precisely, the result is as follows.

COROLLARY: *If C and C' are Cantor sets contained in $E^{(r)}$, $r \geqq 2$, and C is plane or linear then C and C' are equivalent under automorphism of $E^{(r)}$ if and only if some homeomorphism of C onto C' can be extended to a neighborhood of C. If f is any homeomorphism of C onto C' that can be extended to a neighborhood of C then f can be extended to $E^{(r)}$ in such a way that it is equal to the identity outside any given rectangular r-cell to which C and C' are interior.*

Only the second assertion need be proved since it obviously implies the first. We shall consider first the case in which C is linear. Let R be the given rectangular r-cell containing C and C' in its interior and let f_1 be an extension of f to an open set G containing C. We may suppose that both G and f_1G are contained in R. Using the Heine-Borel theorem we can find a finite number of disjoint rectangular r-cells $R_1, \cdots, R_N$ contained in G whose interiors cover C. Let p_i be the center of R_i. Then $f_1p_1, \cdots, f_1p_N$ are distinct interior points of R and by Lemma 13 there exists $h_1 \in H_0[R]$ such that $h_1f_1p_i = p_i$. Let σ_i' be an r-cell about p_i so small that $h_1f_1\sigma_i' \subset R_i$, and let σ_i be an r-cell interior to R_i and containing CR_i in its interior. Let g be a shrinking transformation of the interior of each cell R_i such that $g\sigma_i \subset \sigma_i'$ and such that g leaves boundary points of $R_1, \cdots, R_N$ fixed. Then h_1f_1g is a homeomorphism of each σ_1 onto a subset of R_i. By Lemma 16 applied to each cell R_i there exists $h_2 \in H_0[R]$ such that $h_2h_1f_1g$ is equal to the identity on C. From the theorem on invariance of region it follows as before that the transformation h_0 equal to $f_1gf_1^{-1}$ in $f_1R_1 + \cdots + f_1R_N$ and equal to the identity elsewhere is an automorphism, and it evidently belongs to $H_0[R]$ since $f_1G \subset R$. Hence $h = h_2h_1h_0$ is in $H_0[R]$ and $hfp = h_2h_1f_1gp = p$ for every p in C. Therefore h^{-1} is an extension of f to R, and it may be defined equal to the identity outside R.

It remains to consider the case in which C is plane but not linear. Let R be a rectangular r-cell containing both C and C' in its interior and let C'' be a linear Cantor set in the same plane as C and likewise interior to R. Antoine has shown that there is an automorphism of this plane that carries C into C'' and is equal to the identity outside R. This automorphism can be extended to $E^{(r)}$ in such a way that it is still equal to the identity outside R. This can be done conveniently by projecting the transformation from two points in R on opposite sides of the plane in which C lies, then projecting again from two points in R on opposite sides of the 3-space in which the transformation is already defined. After $r - 2$ steps the desired extension is obtained. Thus we have an automorphism $h \in H_0[R]$ that carries C into C''. The transformation fh^{-1} is a homeomorphism of C'' onto C' that can be extended to a neighborhood of C''. Since C'' is linear there exists an automorphism $g \in H_0[R]$ equal to fh^{-1} on C'', as we have already shown. The automorphism gh is therefore equal to f on C and it can be defined equal to the identity outside R.

LEMMA 17: *If the given homeomorphism f in Lemma 16 is measure-preserving then the automorphism h can also be taken to be measure-preserving.*

Let h be any automorphism fulfilling the requirements of Lemma 16. In R consider the measure function $\mu A = mh^{-1}A$. Then $\mu C = mh^{-1}C = mfC = 0$, since f is now assumed to be measure-preserving. Let L denote a line segment containing C and contained in the interior of R. It is possible that μL may be positive, but in that case we can displace the points of L not in C by an automorphism $g_1 \in H_0[R]$ so chosen that $\mu g_1 L = 0$ and $g_1 p = p$ for $p \in C$. Then consider the measure $\mu_1 A = \mu g_1 A = mh^{-1}g_1 A$. This is an r-dimensional LS measure and in addition $\mu_1 L = 0$. Hence, by Theorem 2 Corollary 4, there exists an automorphism $g_2 \in H_0[R]$ such that $\mu_1 A = mg_2 A$ and $g_2 p = p$ for $p \in L$. Thus we have $mh^{-1}g_1 A = mg_2 A$ for all $A \subset R$, that is, $mg_2 g_1^{-1} hA = mA$. Hence $g_2 g_1^{-1} h$ is measure-preserving, and since both g_1 and g_2 leave the points of C fixed the measure-preserving automorphism $T = g_2 g_1^{-1} h$ fulfills the requirements of the lemma.

7. Proof of Lemma 5

LEMMA 18: *Let R_0 be the interior of a rectangular r-cell R, $r \geqq 2$, and let T be a measure-preserving homeomorphism of an open set $G \subset R_0$ onto an open set $TG \subset R_0$, where $mG = mR$, and let $\sigma_1, \cdots, \sigma_N$ be the cells of any given dyadic subdivision. There exist arbitrarily small automorphisms $h_1, h_2 \in H_0[R]$ such that $h_1 T h_2$ is a measure-preserving homeomorphism of $h_2^{-1}G$ onto $h_1 TG$, where $mh_2^{-1}G = mR$, and such that $h_1 T h_2$ transforms a certain Cantor set C_1 in the following manner: The points of C_1 are all periodic with the same period; the distinct images of C_1 are disjoint and equally distributed among the cells $\sigma_1, \cdots, \sigma_N$; the images of C_1 contain a prescribed fraction $\alpha < 1$ of the measure of each of the cells σ_i.*

In the proof we shall assume $mR = 1$, which involves no loss of generality. Observe first that T may be considered as a measure-preserving automorphism of the set $D = \prod T^n G$, $-\infty < n < +\infty$, which has measure one, since it is the intersection of countably many sets of measure one. By Lemma 15, there exist arbitrarily small automorphisms $S_1, S_2 \in M_0[R]$ such that $T_1 = S_2 S_1 T S_2^{-1}$ permutes cyclically the centers of the cells $\tau_1, \cdots, \tau_{KN}$ of an arbitrarily fine dyadic subdivision, in particular, finer than the given subdivision $\sigma_1, \cdots, \sigma_N$. Now consider T_1 to have domain $S_2 G$, and let $\tau \subset \tau_1$ be an r-cell about the center of τ_1 so small that its first KN images under T_1 are defined and respectively interior to $\tau_2, \tau_3, \cdots, \tau_{KN}, \tau_1$. Let C be a linear Cantor set interior to τ. Since T_1^{KN} is a measure-preserving homeomorphism of τ onto a subset of τ_1, by Lemma 17 there exists an automorphism $S_3 \in M_0[\tau_1]$, which we define equal to the identity outside τ_1, such that under $T_2 = S_3 T_1$ the set C is brought back to coincidence with itself. Under T_2 the points of C are all periodic with period KN, and there is one image of C in each of the cells $\tau_1, \cdots, \tau_{KN}$. Now introduce a measure μ_1 in C by mapping it onto a linear Cantor set with measure $1/KN$, and extend this measure to the images of C by the transformation T_2. Then μ_1 will be a normalized measure invariant under T_2 such that $\mu_1 \tau_i = 1/KN$, $i = 1, \cdots, KN$. Since T_2 preserves both measures μ_1 and m, it will also preserve the measure $\mu A = (1 - \alpha) mA + \alpha \mu_1 A$. It is easily verified that μ^* fulfills the

conditions of Theorem 2_1 with respect to each of the cells τ_i. Hence there exists an automorphism $h \in H_0[R]$, which transforms each cell τ_i into itself, such that $\mu^*A = m^*hA$ for all $A \subset R$. Hence $T_3 = hT_2h^{-1}$ will preserve Lebesgue measure, and the Cantor set $C_1 = hC$ will be periodic under T_3, will have one image in each cell τ_i, and this image will contain the fraction α of the measure of τ_i. Hence the images of C_1 under T_3 are equally distributed among $\sigma_1, \cdots, \sigma_N$, and contain the fraction α of their measure. Expressing T_3 in terms of T, we have $T_3 = hS_3S_2S_1TS_2^{-1}h^{-1}$. The automorphisms S_1 and S_2 could be taken arbitrarily small, and S_3 and h have norms no greater than the diameter of the cells τ_i. Hence T_3 is of the required form h_1Th_2, where h_1 and h_2 belong to $H_0[R]$ and are arbitrarily small. It should be noted that T_3 is measure-preserving even though h_1 and h_2 are not.

Taking $\alpha = \frac{1}{2}$, it is evident that Lemma 18 implies Lemma 5, in fact the periodicity of the set C_1 is for this purpose superfluous.

8. The Borel class of the set of metrically transitive automorphisms

To complete the proof of Theorem 1, it only remains to show that the set of metrically transitive automorphisms is G_δ in $M[E, \mu]$. This is most easily done by utilizing what amounts to a necessary and sufficient condition for metrical transitivity, rather than the definition itself. Consider the space L_2 of functions quadratically integrable over E. Any automorphism $T \in M[E, \mu]$ induces a unitary transformation $Uf(x) = f(Tx)$ of L_2, as is well-known.[37] Let $f_1, f_2, \cdots$ be a sequence of continuous functions dense in L_2. Let $E(i, j, n)$ be the set of all T in $M[E, \mu]$ such that

$$\int_E \left[\frac{1}{n}\sum_{r=0}^{n-1} f_i(T^r x) - (f_i, 1)\right]^2 dx < \frac{1}{j},$$

where $(f_i, 1)$ denotes the number $\int_E f_i(x)\,dx$. This set is open in $M[E, \mu]$. For suppose T_k, $k = 1, 2, \cdots$, belongs to the complement of $E(i, j, n)$ and that $\rho(T_k, T) \to 0$ as $k \to \infty$. The integrand converges boundedly to the limit and so T also belongs to the complement. Now, the set M_T of metrically transitive automorphisms is represented by

$$M_T = \prod_{i=1}^{\infty}\prod_{j=1}^{\infty}\sum_{n=1}^{\infty} E(i, j, n).$$

For if T is metrically transitive it will belong to $E(i, j, n)$ for all sufficiently large n, in virtue of the mean ergodic theorem of von Neumann.[38] On the other hand, if T is metrically intransitive there exists an invariant function φ not constant, namely, the characteristic function of an invariant set with measure inter-

[37] B. O. Koopman, *Hamiltonian systems and transformations in Hilbert space*. Proc. Nat. Acad. USA. **17** (1931) 315.

[38] J. von Neumann, *Proof of the quasiergodic hypothesis*. Proc. Nat. Acad. USA. **18** (1932) 70.

mediate between zero and μE. Let d denote the distance in L_2 from φ to the axis of constant functions, and choose $1/j < \frac{1}{4} d^2$. Choose f_i from the sphere of radius $d/2$ about φ. Since φ is a fixed point under U, and U is unitary, all images of f_i under U also lie in the sphere, and therefore also all averages $\frac{1}{n} \sum_{r=0}^{n-1} f_i(T^r x)$. For no n does T belong to $E(i, j, n)$, because the distance from $\frac{1}{n} \sum_{r=0}^{n-1} f_i(T^r x)$ to the constant function $(f_i, 1)$ is always at least $d/2$, and therefore the squared distance is greater than $1/j$.

This representation of M_T evidently exhibits it as a G_δ set in $M[E, \mu]$.

IV. Metrically Transitive Flows

9. Existence Theorem

The results concerning metrically transitive *automorphisms* contained in the preceding sections make it possible to set up a general procedure for defining metrically transitive *continuous flows*.[39] By a continuous flow we shall mean a one-parameter group of automorphisms T_λ, $-\infty < \lambda < +\infty$, of a space E such that $T_\lambda x$ is continuous in x and λ, and such that $T_{\lambda+\mu} = T_\lambda T_\mu$ and $T_0 = I$.

Theorem 3: *Let E be a regularly connected polyhedron of dimension $r \geqq 3$, and let μ be any r-dimensional Lebesgue-Stieltjes measure in E. There exists a continuous flow in E which is metrically transitive with respect to μ, and which leaves all singular points of E fixed.*

It is sufficient to define a continuous flow in a rectangular r-cell which leaves the boundary fixed and is metrically transitive with respect to ordinary Lebesgue measure. Because the image of such a flow under the map defined in Lemma 3 will be a continuous flow in E, in virtue of Lemma 4, and metrically transitive with respect to μ.

We first define a flow in a space Q_1 defined as follows. Let B denote the $(r-1)$-dimensional unit cube and introduce in it a measure $\nu A = \int_A f(p)\, dp$, where the integral is an ordinary $(r-1)$-dimensional Lebesgue integral and $f(p)$ is a continuous function positive over the interior of B and tending to infinity at the boundary in such a way that $\int_B f(p)\, dp = 1$. Then νA is an $(r-1)$-dimensional LS measure in B, and by Theorem 1 there exists an automorphism T of B which leaves the boundary fixed and is metrically transitive with respect to ν. Consider the product space of B with the unit interval $0 \leqq \lambda \leqq 1$, and identify points $(p, 1)$ and $(Tp, 0)$. In this space Q_1 define a flow upward along streamlines perpendicular to B, taking the velocity at any point to be $1/f(p)$, where p is the last intersection of the streamline with B. The velocity along a streamline undergoes a discontinuous change when it crosses B. Nevertheless,

[39] This is an adaptation of a standard method. See E. Hopf, *Ergodentheorie*, p. 41.

the flow defined by this velocity field is continuous, and since $1/f(p)$ tends to zero at the boundary, the boundary remains fixed. This flow preserves r-dimensional Lebesgue measure in Q_1. To see this, consider any small cube interior to Q_1 with height h and base σ. Until it crosses B, the segments along the streamlines are rigidly translated, so that its volume is not changed. After it crosses B it has a new cross-section $T\sigma$, and the length along any streamline is now $h \cdot f(p)/f(Tp)$. The volume is therefore

$$\int_{T\sigma} h \frac{f(p)}{f(Tp)}\, dp = \int_{T\sigma} \frac{h}{f(Tp)}\, d\nu(p) = \int_{\sigma} \frac{h}{f(p)}\, d\nu(p) = \int_{\sigma} h\, dp.$$

Thus, the change in velocity exactly compensates for the change in cross-section, and therefore the flow is measure-preserving, since a flow that preserves the measure of small cubes preserves the measure of all sets. To see that it is metrically transitive, consider any invariant measurable set. The set must be cylindrical, and so its intersection A with B is Lebesgue measurable and therefore ν-measurable. Since A is invariant under T its ν-measure is either zero or one, and consequently also its Lebesgue measure. Hence the cylindrical set over it has Lebesgue measure zero or one.

The space Q_1 in which we have just defined a flow is homeomorphic to the r-dimensional tube Q, that is, the product space of B with $(0, 1)$ where points $(p, 0)$ and $(p, 1)$ are identified. To prove this, observe that since T leaves the boundary fixed, it can be joined isotopically to the identity. That is, there exists a continuous family of automorphisms T_λ of B, $0 \leqq \lambda \leqq 1$, such that $T_1 = T$ and T_0 is the identity.[40] The correspondence $(p, \lambda) \to (T_\lambda p, \lambda)$ is a homeomorphism of Q onto Q_1.

The r-dimensional unit cube R can be represented as the continuous image of Q under a map which is a homeomorphism up to the boundary of Q. This may be done as follows. The region of r-space defined by the inequalities $1 \leqq (x_1^2 + x_2^2)^{\frac{1}{2}} \leqq 2$, $0 \leqq x_i \leqq 1$, $i > 2$, is evidently homeomorphic to Q. The map defined by $x_1' = x_1[(x_1^2 + x_2^2)^{\frac{1}{2}} - 1]$, $x_2' = x_2[(x_1^2 + x_2^2)^{\frac{1}{2}} - 1]$, $x_i' = x_i$, $i > 2$, is continuous and has for range the set $0 \leqq (x_1'^2 + x_2'^2)^{\frac{1}{2}} \leqq 2$, $0 \leqq x_i' \leqq 1$, $i > 2$, which is homeomorphic to R. This map is 1:1 except at points where $(x_1^2 + x_2^2)^{\frac{1}{2}} = 1$, which belong to the boundary of Q. Combining this map with an automorphism of R based on Theorem 2, the map from Q_1 to R can be made measure-preserving. The image in R of the flow already defined in Q_1 is therefore metrically transitive and leaves the boundary of R fixed. From this we can derive a metrically transitive continuous flow in E, as already explained.

10. Most general polyhedra that can support metrically transitive automorphisms or flows

In the present section we consider arbitrary finite polyhedra, not necessarily connected or even homogeneous-dimensional, and seek to characterize those in

[40] J. W. Alexander, *On the deformation of an n-cell*. Proc. Nat. Acad. USA. **9** (1923) 406–407.

which metrically transitive continuous flows or automorphisms are possible. It will be seen that a complete answer is obtained except in the case $r = 2$ for flows.

THEOREM 4: *A finite polyhedron of dimension $r \geqq 3$ can support a metrically transitive continuous flow with respect to any given r-dimensional Lebesgue-Stieltjes measure if and only if its regular points form a connected set.*

If the regular points form a connected set, the closure of this set is a regularly connected polyhedron, and we have already seen how to construct a metrically transitive flow in this part. Since the flow given by Theorem 3 leaves singular points fixed, it can be extended to the rest of the polyhedron, which is a set of measure zero, by defining it equal to the identity there. On the other hand, if the regular points fall into two or more disjoint open sets, these will be invariant under any continuous flow, since if one point of a streamline is regular, all must be. There will therefore necessarily exist disjoint invariant sets with positive measure, so that no metrically transitive, or even topologically transitive, continuous flow is possible.

The case $r = 2$ is not covered by our present method of construction, and it appears likely that further conditions on the polyhedron are necessary in this case. The case $r = 1$ is trivial, the only possibility is for the polyhedron to be homeomorphic to the circumference of a circle, plus possibly some isolated points, the flow being topologically equivalent to a steady rotation of the circle. The details are left to the reader.

THEOREM 5: *A finite polyhedron of dimension $r \geqq 2$ can support a metrically transitive automorphism with respect to any given r-dimensional Lebesgue-Stieltjes measure if and only if its regular components have equal measure and can be permuted cyclically by an automorphism.*

Suppose the conditions satisfied. Let h be an automorphism that permutes the regular components cyclically. By Theorem 2 Corollary 3, we can modify h within each regular component so as to make it measure-preserving. Call the resulting automorphism T. Suppose there are n regular components, and let E_1 be one of them. Then T^n is a measure-preserving automorphism of E_1. By the Corollary to Theorem 1, there exists an arbitrarily small automorphism S of E_1 such that ST^n is a metrically transitive automorphism of E_1, and since S leaves all singular points fixed, it can be extended to the rest of E by defining it equal to the identity outside E_1. Then ST is a metrically transitive automorphism of E, because any invariant set of positive measure must contain almost all points of E_1 and therefore almost all points of every regular component. Conversely, if there exists a metrically transitive automorphism it must permute the regular components cyclically, and consequently they must have equal measure. In fact, these conditions are necessary if there is to exist even a topologically transitive measure-preserving automorphism.

It may be added that in the case $r = 1$ the only polyhedron that can support a metrically transitive automorphism is one that is homeomorphic to a finite number of disjoint circumferences of circles of equal measure, plus possibly some

isolated points. In such a polyhedron a metrically transitive automorphism is easily defined by composing a non-periodic rotation of one of the circles with a cyclic permutation of the circles. The isolated points, if any, may be left fixed. Any other 1-dimensional polyhedron will have a regular component with at least one singular point. This component will be either a line segment or a circle with one singular point. Some power of any given automorphism will leave all singular points fixed and each regular component invariant, since there are only a finite number of each. But a measure-preserving automorphism of a line segment that leaves the ends fixed must be the identity, and a measure-preserving automorphism of a circle that leaves one point fixed must be either the identity or a reflection in a diameter. In none of these cases can the automorphism be metrically transitive, since a power of it will be equal to the identity on some regular component.

Thus we have obtained a characterization of all finite polyhedra that can support metrically transitive automorphisms.

It may be remarked that although metrically transitive automorphisms may exist in polyhedra having more than one regular component, nevertheless Theorem 1 is not true for such polyhedra, because we have just seen that a transitive automorphism must permute the regular components cyclically, and therefore cannot be near the identity. However, the following generalization of Theorem 1 holds. *In the space $M[E, \mu]$ of measure-preserving automorphisms of any polyhedron that can support at least one metrically transitive automorphism, the metrically transitive automorphisms form a residual set with respect to the subspace of automorphisms that permute the regular components cyclically.* One need only observe that in the existence proof given above it was really shown that the metrically transitive automorphisms are dense in this subspace. Since they form a G_δ set with respect to the whole space, they are G_δ with respect to the subspace also, and therefore residual.

11. Transitive automorphisms in transitive flows

In the preceding sections we have made use of a general method whereby a transformation can be used to define a flow in a product space. By this construction a transitive transformation gives rise to a transitive flow. But there is an even simpler relation connecting transformations and flows, namely, the individual transformations that make up a continuous flow are themselves automorphisms of the space. If even one automorphism in the flow is transitive, either metrically or topologically, then the flow is, because by definition a set that is invariant under a flow is invariant under each of the transformations that make up the flow. This might seem to suggest another way of deriving a transitive flow from a transformation, but the method is of little use because in general an automorphism cannot be embedded in a flow.[41] Nevertheless, it is natural

[41] Any transformation embedded in a continuous flow must, for example, have roots of all orders. A simple example of an automorphism which hasn't even a square root is the following. Let T_1 be a rotation of the circumference of a circle through an angle π/k,

to inquire whether any of the automorphisms that make up a transitive continuous flow are necessarily transitive. In the case of metrical transitivity this question appears to be open.[42] It is therefore of interest to note that in the topologically transitive case the answer is in the affirmative, as we proceed to show.

THEOREM 6: *Let T_λ, $-\infty < \lambda < +\infty$, be a topologically transitive continuous flow in a separable metric space E, and suppose there is no isolated streamline. For all values of λ, except a set of first category on the line $-\infty < \lambda < +\infty$, the automorphisms T_λ are topologically transitive.*

Let $G_1, G_2, \cdots$ be a countable base for all open sets in E. Let $A_{i,j}$ be the set of values of λ such that for some positive or negative integer k the set $T_{k\lambda}G_i$ overlaps G_j, that is, such that G_i overlaps G_j under some power of T_λ. The set $A_{i,j}$ is evidently open, we proceed to show that it is dense on the line $-\infty < \lambda < +\infty$. Consider any interval I and form the set of all numbers $k\lambda$, where $\lambda \epsilon I$ and $k = 0, \pm 1, \pm 2, \cdots$. This set includes all numbers greater in absolute value than some number Λ. Form the set $H = \sum_{|\lambda|>\Lambda} T_\lambda G_i$ and consider a sphere $\sigma \subset G$ so small that the set swept out by it as λ describes the interval $-\Lambda \leqq \lambda \leqq \Lambda$ is not dense in G_j. Any sufficiently small sphere about a point of G_i will do since by hypothesis there is no isolated streamline. The set H must overlap G_j, because otherwise the invariant open set swept out by σ would not be dense in G_j, contrary to the hypothesis that the flow is topologically transitive. Hence there exist values of λ greater in absolute value than Λ such that $T_\lambda G_i$ overlaps G_j, and so the set $A_{i,j}$ contains points of the interval I. Since the sets $A_{i,j}$ are open and everywhere dense, their intersection as i and j run independently over all positive integers is a residual set. For any value of λ in this set, the automorphism T_λ is topologically transitive, because if it had an invariant open set not everywhere dense in E there would exist a pair G_i, G_j such that no image of G_i overlaps G_j, and this would contradict the fact that λ belongs to $A_{i,j}$.

V. SOME RELATED QUESTIONS

12. Generation of Random Sequences by Automorphisms

In the Introduction it was explained that metrical transitivity implies that the images of almost all points are distributed like random sequences in respect

---

where k is a positive integer. Let T_2 be the automorphism of the segment $0 \leqq \theta \leqq \frac{\pi}{k}$ defined by $T_2\theta = \left(\frac{k\theta}{\pi}\right)^2$, and let it leave all other points fixed. Under the automorphism $T = T_2T_1$, the points $0, \frac{\pi}{k}, \frac{2\pi}{k}, \cdots, \frac{(2k-1)\pi}{k}$ form a single cycle of period $2k$, and no other points are periodic. It is clear that T cannot be the square of a transformation S, because a cycle of period $2k$ in T could arise only by the splitting of a cycle of period $4k$ in S, and then T would have two cycles of period $2k$. Hence T has no square root.

[42] The question was raised by H. E. Robbins.

to the frequency with which they fall in any given measurable set. A random sequence has also the property that of its first n points the number that fall in A differs from $n \cdot mA$ by something of the order of $\sqrt{n}$. Furthermore, this difference oscillates in sign. Let us consider the r-dimensional unit cube R and take for A any dyadic cell σ. Let the characteristic function of σ be f. The difference in question is then $\sum_{\nu=1}^{n} f(T^\nu p) - nm\sigma$. Consider instead the difference $K \cdot \sum_{\nu=1}^{n} f(T^\nu p) - n$, where $K = 1/m\sigma$ is the number of cells in the dyadic subdivision to which σ belongs. This latter difference is an integer, and as n increases by unity the difference either increases by $K - 1$ or decreases by 1. To pass from a positive to a negative value it must therefore pass through the value zero. For a random sequence this must happen infinitely often. This means that a random sequence from R has arbitrarily long segments distributed between σ and its complement in *exact* proportion to their measures. We proceed to show that there exist automorphisms of R under which almost all points generate sequences that share this property of random sequences with respect to every dyadic cell, and indeed that such automorphisms are the "general case."

THEOREM 7: *Let R denote the r-dimensional unit cube, $r \geqq 2$. There is a residual set of automorphisms in $M[R]$ under which almost all points p generate sequences distributed in the following manner: Given any dyadic subdivision, there exist arbitrarily large values n such that the first n images of p are equally distributed among the cells of this subdivision.*

Let $\sigma_1^{(j)}, \cdots, \sigma_{N_j}^{(j)}$ be the cells of the j-th dyadic subdivision, and let $f_i^{(j)}$ be the characteristic function of $\sigma_i^{(j)}$. Let $E(i, j, k, n)$ be the set of all T in $M[R]$ such that the measure of the set $A(i, j, n, T)$ of points p for which $\frac{1}{n}\sum_{\nu=1}^{n} f_i^{(j)}(T^\nu p) = m\sigma_i^{(j)}$ is greater than $1 - \frac{1}{k}$. The set $E(i, j, k, n)$ is open, because if T belongs to it and G_ϵ denotes the ϵ-neighborhood of the boundary of $\sigma_i^{(j)}$, where ϵ is so chosen that $mG_\epsilon < \frac{1}{n}\left[mA(i, j, n, T) - \left(1 - \frac{1}{k}\right)\right]$, we have $m\left[A(i, j, n, T) - \sum_{\nu=1}^{n} T^{-\nu}G_\epsilon\right] > 1 - \frac{1}{k}$. If p belongs to $A(i, j, n, T) - \sum_{\nu=1}^{n} T^{-\nu}G_\epsilon$, then $Tp, T^2p, \cdots, T^np$ all lie outside G_ϵ, and so $\sum_{\nu=1}^{n} f_i^{(j)}(T_1^\nu p) = \sum_{=1}^{n} f_i^{(j)}(T^\nu p)$ for any automorphism T_1 so near to T that $\rho(T_1^\nu, T^\nu)$ is less than ϵ for $\nu = 1, \cdots, n$. Consequently, the set $A(i, j, n, T_1)$ contains $A(i, j, n, T) - \sum_{=1}^{n} T^{-\nu}G_\epsilon$, and therefore has measure exceeding $1 - \frac{1}{k}$, so that T_1 belongs to $E(i, j, k, n)$. Hence $\sum_{n=N}^{\infty} \prod_{i=1}^{N_j} E(i, j, k, n)$ is likewise open. But this is the set of automorphisms such

that for some $n \geqq N$ the set of points whose images are distributed equally among $\sigma_1^{(j)}, \cdots, \sigma_{N_j}^{(j)}$ has measure greater than $1 - \frac{1}{k}$. Lemma 18 implies that such automorphisms are everywhere dense in $M[R]$, as may be seen by taking $\alpha > 1 - \frac{1}{k}$, because an automorphism that has a periodic set whose images are equally distributed among $\sigma_1^{(j)}, \cdots, \sigma_{N_j}^{(j)}$ and have combined measure α will belong to $E(i, j, k, n)$ for any value of n divisible by the period of the set. Therefore the set $\prod_{j=1}^{\infty} \prod_{N=1}^{\infty} \prod_{k=1}^{\infty} \sum_{n=N}^{\infty} \prod_{i=1}^{N_j} E(i, j, k, n)$ is residual, since it is an intersection of countably many dense open sets. If T belongs to this set, then, for every pair j, N, almost all points will generate sequences which have segments of length greater than N that are equally distributed among the cells of the j-th dyadic subdivision. The intersection of these sets as j and N run over all positive integers will still have measure one, and the points of this set will have images distributed in the required manner.

13. Automorphisms that preserve sets of measure zero

We shall say that an automorphism *preserves zero sets* if it has the property that $mhA = 0$ if and only if $mA = 0$. This is equivalent to requiring that hA be measurable if and only if A is measurable, because any Lebesgue measurable set is the sum of a Borel set and a set of measure zero, and therefore its image is the sum of a Borel set and the image of a zero set. Consequently, if h preserves zero sets it will also preserve measurable sets. On the other hand, if it does not preserve zero sets either h or its inverse must take some zero set into a set with positive outer measure. Therefore it takes some zero set into a non-measurable set, because every set with positive outer measure contains a non-measurable subset.[43] Hence automorphisms that preserve zero sets might equally well be called *measurability preserving*. It is well-known that an automorphism need not preserve zero sets, but we proceed to show that any automorphism is topologically equivalent to one that does.

THEOREM 8: *Let h be any automorphism of a rectangular r-cell R (or of $E^{(r)}$), $r \geqq 1$. There exists an automorphism g of R (or $E^{(r)}$) such that ghg^{-1} preserves zero sets.*

Consider first the case of an automorphism of R. Define $\mu^*A = \frac{1}{2}m^*A + \sum_{n=1}^{\infty} \frac{1}{2^{n+2}} (m^*h^nA + m^*h^{-n}A)$. Then μ is an r-dimensional LS measure in R, and $\mu R = mR$. By Theorem 2 there exists an automorphism g such that $\mu A = mgA$. Evidently $\mu A = 0$ if and only if $mh^nA = 0$ for every positive or negative integer n. Therefore $\mu hA = 0$ if and only if $\mu A = 0$; that is, $mghg^{-1}A = 0$ if and only if $mA = 0$.

Now consider any automorphism h of $E^{(r)}$. Divide the space into a lattice

[43] Carathéodory, *Vorlesungen über reelle Funktionen*, Leipzig and Berlin 1918, p. 354.

of cubes $R_1, R_2, \cdots$ and let G_i be the interior of R_i. Form the outer measure

$$\mu^*A = \sum_{i=1}^{\infty} \left\{ \frac{1}{2} \cdot \frac{m^*AG_i}{mG_i} + \sum_{n=1}^{\infty} \frac{1}{2^{n+2}} \left[\frac{m^*h^n(AG_i)}{mh^nG_i} + \frac{m^*h^{-n}(AG_i)}{mh^{-n}G_i} \right] \right\}.$$

This measure is evidently zero for points and positive for non-void open sets. Furthermore, $\mu R_i = 1$, and therefore bounded sets have finite measure, while $\mu E^{(r)} = +\infty$. Hence, by Theorem 2 Corollary 5, there exists an automorphism g such that $\mu^*A = m^*gA$, at least in case $r \geqq 2$. The same argument as above then shows that ghg^{-1} preserves zero sets. In the case $r = 1$, the positive and negative half lines both have infinite μ-measure, so that the proof is valid in this case also.

14. Sets automorphic to zero sets

We shall say that two subsets A and B of $E^{(r)}$ are automorphic if there exists an automorphism h of $E^{(r)}$ such that $hA = B$. This is evidently a stronger notion than that of homeomorphism of the two sets. We shall obtain a simple characterization of sets that are automorphic to sets of measure zero.

THEOREM 9: *Let B be any subset of $E^{(r)}$, $r \geqq 1$. In order that there exist an automorphism h of $E^{(r)}$ such that hB has Lebesgue measure zero it is necessary and sufficient that the complement of B contain a sequence of perfect sets whose union is dense in $E^{(r)}$. If a bounded set B satisfies this condition, the automorphism h can be taken equal to the identity outside of any cube that contains B.*

The condition is evidently necessary, since if $mhB = 0$ the complement of hB contains a sequence of perfect sets P_n whose union contains all the measure in $E^{(r)}$, and is therefore dense. The perfect sets $h^{-1}P_n$ are therefore outside B and their union is also dense in $E^{(r)}$. To prove that the condition is also sufficient, suppose that B satisfies it and consider any cube R with interior R_0. By hypothesis $R_0 - R_0B$ contains a dense sequence of perfect sets. In $R_0 - R_0B$ it is therefore possible to find a sequence of perfect sets P_n such that every neighborhood in R contains at least one member of the sequence. In each set P_n there exists a Cantor set C_n. Introduce a normalized LS measure μ_n in C_n by mapping it homeomorphically onto a linear Cantor set with linear measure one. Form the outer measure $\mu^*A = mR \cdot \sum_{n=1}^{\infty} \frac{1}{2^n} \mu_n^* AC_n$. This is evidently a LS outer measure in R, zero for points and positive for non-void open sets in R. Also $\mu R = mR$ and $\mu BdR = 0$. By Theorem 2_1 there exists an automorphism $h \in H_0[R]$ such that $\mu^*A = m^*hA$. Since $\mu BR = 0$, we have $mh(BR) = 0$. If B is contained in R we can take h equal to the identity outside R. If B is an unbounded set we can divide $E^{(r)}$ into a lattice of cubes and transform each into itself in such a way as to compress the part of B contained in it to a set of measure zero. The resulting automorphism of $E^{(r)}$ will then carry B into a zero set.

Since any residual set in $E^{(r)}$ contains a dense sequence of perfect sets, it follows as a corollary that any set of first category is automorphic to a zero set.

In a previous paper[44] we have obtained this result directly by a simple category argument. In the case $B \subset R$, we showed in fact that the automorphisms h such that $mhB = 0$ form a residual set in $H[R]$. It is interesting to note that it is *only* for first category sets that this method of proof is available, at least for sets having the property of Baire, because in the same paper it was shown (Theorem 4) that a set having the property of Baire is of first or second category according as the automorphisms that carry it into a set of measure zero form a residual set or a set of first category in $H[R]$. The present theorem may be regarded as completing the earlier result by showing precisely which second category sets are automorphic to zero sets.

15. Transformations topologically equivalent to measure-preserving automorphisms

G. D. Birkhoff has formulated the following problem: Given an automorphism h, when can one assert that there exists an automorphism g such that ghg^{-1} is Lebesgue measure-preserving? In other words, the problem is to characterize *in topological terms* automorphisms that are topologically equivalent to measure-preserving ones. It is well-known that measure-preserving automorphisms have special topological properties, such as the recurrence property discovered by Poincaré and similar topological properties that can be deduced from the ergodic theorem. Again, it is obvious that a measure-preserving automorphism cannot transform a closed sphere into a subset of its interior. Thus it is clear that h must be restricted, in contrast to what we found in the case of the related problem of characterizing automorphisms equivalent to ones that preserve zero sets (§13). In the present section we shall obtain a result that constitutes a solution of this problem, though it must be admitted that the characterization obtained is neither particularly simple nor easy to apply. It is to be hoped that a more elegant solution may be found.

Let us restrict attention to an r-dimensional cube R. If there exists an automorphism g such that ghg^{-1} preserves Lebesgue measure, then h preserves the measure $\mu A = mgA$; and conversely, if h preserves a measure μ automorphic to Lebesgue measure, there will exist a g such that ghg^{-1} preserves Lebesgue measure, namely, any automorphism such that $\mu A = mgA$. Thus Theorem 2 provides the following equivalent formulation of the problem: Given an automorphism h, when can one assert that it preserves some r-dimensional LS measure? It is from this point of view that we shall approach the problem. In a previous paper[45] we have discussed the question of the existence of somewhat more general invariant measures. We obtained the result (loc. cit. Th. 2) that an automorphism T of a complete separable metric space admits a finite invariant Borel measure that is zero for points if (i) T has non-denumerably many

[44] See footnote 28.

[45] J. C. Oxtoby and S. M. Ulam, *On the existence of a measure invariant under a transformation*. Ann. of Math. **40** (1939) 560–566.

periodic points, or (ii) there exists a compact set C consisting of non-periodic points of which at least one returns to C with positive frequency under iteration of T. (A point p of C is said to return with positive frequency if $\lim_{n\to\infty} \frac{1}{n} \sum_{\nu=1}^{n} f_C(T^\nu p) > 0$, where f_C is the characteristic function of C.) It should be added that in case (i) the invariant measure was defined in such a way that any prescribed sphere containing non-denumerably many periodic points could be made to have positive measure, and in case (ii) positive measure was ascribed to the set C. Assuming this result, we can obtain the following theorem.

THEOREM 10: *Let h be an automorphism of the r-dimensional unit cube R, $r \geqq 2$. In order that there exist an automorphism g such that ghg^{-1} preserves Lebesgue measure it is necessary and sufficient that in every sphere there exist a perfect set to which some point returns with positive frequency under iteration of h, and that these perfect sets can all be chosen outside of an arbitrarily prescribed countable set.*

The necessity of the condition may be deduced from the ergodic theorem as follows. Suppose h is measure-preserving and let A_0 be any given countable set. Then h is a measure-preserving automorphism of the set $E = R - \sum_{n=-\infty}^{+\infty} h^n A_0$, which has measure one. In any sphere there exists a perfect set $P \subset E$ such that $mP > 0$. Denote its characteristic function by $f(p)$. According to the ergodic theorem $f^*(p) = \lim_{n\to\infty} \frac{1}{n} \sum_{\nu=1}^{n} f(T^\nu p)$ exists for almost all p, and $\int f^*(p)\, dp = mP$. Hence $f^*(p)$ is positive for some point p, that is, p returns to P with positive frequency. Thus h has the required property, and therefore also any equivalent automorphism ghg^{-1}, since the condition is topologically invariant.

Conversely, suppose h fulfills the condition. Let $\sigma_1, \sigma_2, \cdots$ be an enumeration of all dyadic cells in R. Let A_n denote the set of periodic points contained in σ_n provided there are only countably many, otherwise let A_n be void. Then the union A_0 of the sets A_n is countable. Consider any cell σ_n. If it contains non-denumerably many periodic points there exists an invariant measure μ_n which is zero for points and positive for σ_n, in virtue of the theorem quoted above. If σ_n contains only countably many periodic points, then by hypothesis there exists a perfect set $P \subset \sigma_n - A_0$ to which some point returns with positive frequency. Since P consists of non-periodic points, the theorem quoted asserts again the existence of an invariant measure μ_n. Thus we obtain a sequence of invariant Borel measures μ_n, each zero for points and such that $\mu_n\sigma_n > 0$. We suppose them normalized. Then $\mu A = \sum_{n=1}^{\infty} \frac{1}{2^n} \mu_n A$ is a normalized invariant Borel measure in R that is zero for points and positive for non-void open sets. To obtain a measure which in addition is zero for the boundary it is sufficient to form $\nu A = \mu A R_0 / \mu R_0$. The Borel measure ν admits an r-dimensional LS extension defined by $\nu^* A = \inf \nu G$, G open, $G \supset A$, which is invariant under h.

By Theorem 2_1 there exists an automorphism g such that $\nu^*A = m^*gA$, and therefore ghg^{-1} preserves Lebesgue measure, as already remarked.

16. A topological ergodic theorem

The paper[45] referred to in the last section also contains the result (Th. 1) that an automorphism T of a complete separable metric space admits a finite invariant Borel measure if there exists a compact set C to which some point returns with positive *superior* frequency $\left(\text{that is, } \limsup_{n\to\infty} \frac{1}{n}\sum_{\nu=1}^{n} f_C(T^\nu p) > 0\right)$. Again it should be added that the measure is positive for C. As a corollary of this, we may assert the following theorem, which properly belongs to the earlier paper since it involves no other results.

THEOREM 11: *Let T be an automorphism of a separable metric space E, and let C be any compact set contained in E. There exists at least one point p in C such that $\lim_{n\to\infty} \frac{1}{n}\sum_{\nu=1}^{n} f_C(T^\nu p)$ exists, where f_C is the characteristic function of C.*

The interest of this theorem is that, like the ergodic theorem, it asserts the existence of a limiting frequency of return for certain points, but *without any measure-theoretic assumptions.* Nevertheless, the proof requires a very considerable excursion into measure theory. It would be interesting to know whether the result can be obtained directly, even in the special case in which C is a square and T an automorphism of the plane.

To prove the theorem, observe that either $\lim_{n\to\infty} \frac{1}{n}\sum_{\nu=1}^{n} f_C(T^\nu p) = 0$ for every point of C, in which case the theorem is true, or there exists a point p in C such that $\limsup_{n\to\infty} \frac{1}{n}\sum_{\nu=1}^{n} f_C(T^\nu p) > 0$. In the latter case there exists a finite invariant Borel measure μ such that $\mu C > 0$, by the theorem quoted above. The ergodic theorem is applicable to this measure and asserts that the limit in question exists for almost all points (in this measure) and is positive for at least one point, since its integral is equal to μC. Therefore there exists a point of C, not necessarily equal to p, which returns with positive frequency.

17. Connections with Haar measure

Haar's Theorem[46] asserts that in any locally compact separable topological group there exists a left (or right) invariant measure which is finite and positive for every compact neighborhood. It is evidently zero for points, and it is unique up to a constant multiplier.[47] In the compact case it is finite, and then left and

---

[46] A. Haar, *Der Massbegriff in der Theorie der kontinuierlichen Gruppen.* Ann. of Math. **34** (1933) 147–169.

[47] J. von Neumann, *The uniqueness of Haar's measure.* Recueil Mathématique, Moscou, N.S. **1** (1936) 721–734.

right invariant measures coincide and are also inverse invariant.[48] The measure is derived from a Carathéodory outer measure that satisfies condition M5.[49] In case the group manifold is a polyhedron, the Haar measure is therefore r-dimensional, and Theorem 2 Corollary 1 establishes the following connection with Lebesgue measure. *If the group manifold of a topological group is a finite euclidean polyhedron, the Haar measure, suitably normalized, is automorphic to the Lebesgue measure in the polyhedron.* Again, we may say that *if two groups have for group manifold the same polyhedron, their Haar measures are automorphic to each other, provided only that they are similarly normalized.*

18. Approximation by locally linear automorphisms

In the proof of Theorem 1, methods were developed that made it possible to approximate any measure-preserving automorphism by one that has a periodic Cantor set whose images are distributed in a regular manner. It is natural to ask whether one can find similarly an approximating automorphism that has a periodic cube. In the present section we shall show that this is possible for automorphisms of a cube that leave the boundary fixed, indeed that an approximating automorphism can be found that has a sequence of periodic cubes which together include almost all points and which the transformation permutes among themselves as if by translation.

It may be remarked that the same methods can be extended to apply to r-dimensional polyhedra embedded in r-space, and to automorphisms that are merely isotopic to the identity. An attempt to extend them to general automorphisms and polyhedra seems to encounter more serious difficulties, and for this reason the proof of Theorem 1 was based on periodic Cantor sets rather than on periodic cubes, which might have seemed more natural.

LEMMA 19: *Let R be the r-dimensional unit cube, $r \geqq 2$, and let T be a measure-preserving automorphism of R that has an invariant set consisting of a finite number of disjoint cubes $R_1, \cdots, R_M$ interior to R and each a sum of cubes of some dyadic subdivision. There exists an arbitrarily small automorphism $S \in M_0[R]$, equal to the identity on $R_1, \cdots, R_M$, such that ST permutes cyclically the centers of the cubes of a certain arbitrarily fine dyadic subdivision of R that are not contained in $R_1, \cdots, R_M$.*

Let E be the polyhedron obtained from R by removing the interiors of $R_1, \cdots, R_M$. Then E is regularly connected, and T may be considered as an automorphism of E. By the Corollary to Theorem 1, T can be made metrically transitive by composing it with an arbitrarily small automorphism that leaves the boundary of E fixed. We may therefore assume without loss of generality that T itself is metrically transitive with respect to E. By Lemma 3, we can represent E as the continuous image of a rectangular r-cell R^* under a measure-

[48] J. von Neumann, *Zum Haarschen Mass in topologischen Gruppen*. Compositio Mathematica **1** (1934) 106–114.

[49] See the note of S. Banach in Saks, *Theory of the Integral*, Appendix 1.

preserving map f which is a homeomorphism up to the boundary. Let ϵ be a given positive number and let $\delta > 0$ be such that any automorphism in $H_0[R^*]$ with norm less than δ corresponds to an automorphism of E with norm less than ϵ. Let $\sigma_1, \cdots, \sigma_N$ be the cubes in E belonging to a dyadic subdivision of diameter less than ϵ. Order these in a sequence $\sigma_{i_1}, \cdots, \sigma_{i_K}$ such that each cube σ_i appears at least once and such that σ_{i_k} and $\sigma_{i_{k+1}}$ have a regular face in common, $1 \leqq k \leqq K-1$. (It may be possible to find such an ordering without repetitions, but this is not necessary.) Since T is metrically transitive on E there exists a non-periodic point whose images under iteration of T fall in the interiors of each of the cubes σ_i with frequency $1/N$, in fact, such points form a set of measure mE. Let p be such a point, then for any sufficiently large n the number of points $Tp, \cdots, T^n p$ in any cube σ_i will differ from n/N by an arbitrarily small fraction of n, in particular by less than n/NK^2. In addition, let us choose n so that $n > NK^2\lambda$ and so that $n + \lambda$ is a number of the form $N \cdot 2^{ar}$, where λ is the integer corresponding to δ defined in Lemma 14. As in the proof of Lemma 15, we can modify the transformation $f^{-1}Tf$ by composition with an automorphism $S_0 \epsilon M_0[R^*]$ such that under $S_0 f^{-1}Tf$ the point $f^{-1}p$ has period $n' = n + \lambda$ and its first n images are the same as under $f^{-1}Tf$. (Let D be the set on which $T_0 = f^{-1}Tf$ is 1:1, and let $p_i^* = f^{-1}T^i p$, $0 \leqq i \leqq n$. Then $p_n^* \epsilon (T_0^{-1}U_\delta)p_0^*$ and step by step we can choose points $p_{n+1}^*, \cdots, p_{n+\lambda-1}^*$ such that $p_0^*, \cdots, p_{n+\lambda-1}^*$ are distinct, and therefore also $T_0 p_0^*, \cdots, T_0 p_{n+\lambda-1}^*$, and such that $T_0 p_i^*$ is always within distance δ of p_{i+1}^*, $p_{n+\lambda}^*$ being taken equal to p_0^*. Then define S_0 with norm less than δ so that $S_0 T_0 p_i^* = p_{i+1}^*$, $0 \leqq i \leqq n + \lambda - 1$.) Then $S_1 = fS_0 f^{-1}$ will be an automorphism of E with norm less than ϵ that leaves the boundary fixed, and under $S_1 T$ the point p will have period n' and its first n images will still be $Tp, \cdots, T^n p$. It follows that the number of distinct images of p under $S_1 T$ in any cube σ_i differs from n'/N by less than $\eta = 2n'/NK^2$, because if σ_i contains n_i of the points $Tp, \cdots, T^n p$, and λ_i of the points $T_1^{n+1}p, \cdots, T_1^{n+\lambda}p$, we have

$$\left| n_i + \lambda_i - \frac{n'}{N} \right| \leqq \left| n_i - \frac{n}{N} \right| + \left| \lambda_i - \frac{\lambda}{N} \right| \leqq \frac{n}{NK^2} + \lambda < \frac{2n}{NK^2} < \eta.$$

Let $p_1, \cdots, p_{n'}$ denote the n' distinct images of p under $S_1 T$. We proceed to assign each of these to one of the cubes σ_i in such a way that exactly n'/N are assigned to each cube σ_i, and each point is assigned either to the cube in which it lies or to an adjacent one. To do this we proceed as follows. First assign each point tentatively to the cube in which it lies. If the number of points in σ_{i_1} is less than n'/N, make up the deficiency by assigning to it points from σ_{i_2}. If the number in σ_{i_1} is greater than n'/N, assign enough points from it to σ_{i_2} to remove the excess. Next consider σ_{i_2} and eliminate the excess or deficiency of points now assigned to it by assigning points of it to σ_{i_3}, or to it from σ_{i_3}. After the k-th step there will be exactly n'/N points assigned to each of the cubes $\sigma_{i_1}, \sigma_{i_2}, \cdots, \sigma_{i_k}$ except those (if any) which are equal to $\sigma_{i_{k+1}}$. After the $(K-1)$-st step there will be exactly n'/N points assigned to each cell

σ_i. At the first step the number of points assigned to other cubes is at most η, at the second step at most 2η are assigned, and at the last step at most $(K-1)\eta$. The total number of points assigned to other cubes is therefore at most $\frac{1}{2}K(K-1)\eta$, which is less than the number of points p_j in any one of the cubes σ_i. Hence the assignments can be made in such a way that no point is assigned to another cube more than once, and therefore every point is assigned finally either to the cube in which it lies or to an adjacent one. We can therefore number the centers $q_1, \cdots, q_{n'}$ of the cubes of the α-th dyadic subdivision of $\sigma_1, \cdots, \sigma_N$ in such a way that the points q_j and p_j always lie in the same cube σ_i or in cubes having an $(r-1)$-face in common. The distance between them is therefore less than 2ϵ, and we can define a measure-preserving automorphism S_2 of E that leaves the boundary fixed, has norm less than 2ϵ, and carries each p_j into q_j. (This does not follow directly from Lemma 13 as stated, but the same method of proof used there evidently applies in the present case since the line segment joining p_j to q_j lies in the interior of E.) It follows that the points $q_1, \cdots, q_{n'}$ are permuted cyclically by $S_2S_1TS_2^{-1}$, which may be written in the form ST by letting $S = S_2S_1TS_2^{-1}T^{-1}$. This completes the proof, since S leaves the entire boundary of E fixed and can be made arbitrarily small by suitable choice of ϵ, since S_1 and S_2 have norms less than 2ϵ.

LEMMA 20: *Let R be the r-dimensional unit cube, $r \geqq 2$, and let $T \in M_0[R]$ be such that it permutes cyclically the centers of some of the cubes of a certain dyadic subdivision of R. Let $\sigma_1, \cdots, \sigma_N$ be the cubes whose centers are permuted. There exist automorphisms $S \in M_0[R]$ and $h \in H_0[R]$, equal to the identity outside the cubes σ_i and on their boundaries, such that $hSTh^{-1}$ is measure-preserving and under it the cubes concentric[50] to $\sigma_1, \cdots, \sigma_N$ with half their diameter are rigidly permuted in a single cycle.*

Let p_i be the center of σ_i. We may suppose the cubes so numbered that $Tp_i = p_{i+1}$, $i < N$, and $Tp_N = p_1$. Let $\tau_1, \cdots, \tau_N$ be cubes concentric to $\sigma_1, \cdots, \sigma_N$, all with the same diameter, and small enough so that $T\tau_i$ is interior to σ_{i+1}, $i < N$, and $T\tau_N$ to σ_1. Let g_1 be an automorphism of R that leaves the boundary fixed and translates τ_1 into τ_2. Extend both g_1 and T to the whole space $E^{(r)}$ by defining them equal to the identity outside R. Let g_2 be a radial shrinking transformation of $E^{(r)}$ such that $g_2R \subset \tau_2$ and such that g_2 leaves both τ_2 and $T\tau_1$ fixed. It is easily verified that $(g_2g_1T^{-1}g_2^{-1})T$ translates τ_1 into τ_2, and $g_2g_1T^{-1}g_2^{-1}$ is an automorphism equal to the identity outside τ_2 that carries $T\tau_1$ into τ_2 in a measure-preserving manner. By Theorem 2 Corollary 3, we can modify $g_2g_1T^{-1}g_2^{-1}$ in $\tau_2 - T\tau_1$ in such a way that it becomes a measure-preserving automorphism of τ_2 that leaves the boundary fixed. Call this modified automorphism S_2 and define it equal to the identity outside τ_2. Then S_2T carries τ_1 rigidly into τ_2. In exactly the same way we can define $S_3 \in M_0[\tau_3]$ such that S_3T carries τ_2 rigidly into τ_3. Finally we define $S_1 \in M_0[\tau_1]$

[50] We shall use the term "concentric" to mean that the cubes have the same center and also have corresponding faces parallel.

such that S_1T carries τ_N rigidly into τ_1. The product of these N transformations S_i defines an automorphism S of R that carries each cube σ_i into itself and is equal to the identity on their boundaries and outside them. Under ST each cube τ_i is translated into τ_{i+1}, $i < N$, and τ_N into τ_1. Finally, let h be a radial transformation of each cube σ_i into itself which carries τ_i into the concentric cube σ_i^* with diameter half that of σ_i. Then $hSTh^{-1}$ is a measure-preserving automorphism that permutes the cubes σ_i^* rigidly and cyclically.

LEMMA 21: *Let T be a measure-preserving automorphism of the r-dimensional unit cube R, $r \geqq 2$, and suppose that T permutes rigidly certain disjoint cubes $R_1, \cdots, R_K$ contained in the interior of R which are sums of cubes of a dyadic subdivision. There exists an arbitrarily small automorphism $S \in M_0[R]$ equal to the identity on $R_1, \cdots, R_K$ such that ST permutes rigidly in a single cycle the cubes concentric to and with half the diameter of those of a certain dyadic subdivision of the complementary set $E = R - (R_1 + \cdots + R_K)$.*

We shall consider the degenerate case $K = 0$ to be included in this statement.

Let $\epsilon > 0$ be given. By the continuity of the group product, there exists a positive number δ such that if h, S_1, S_2 are any automorphisms with norm less than δ then $hS_2S_1Th^{-1}T^{-1}$ will have norm less than ϵ. By Lemma 19 (or by Lemma 15 in the case $K = 0$), there exists $S_1 \in M_0[R]$ equal to the identity on $R_1, \cdots, R_K$ with norm less than δ such that S_1T permutes cyclically the centers of the cubes $\sigma_1, \cdots, \sigma_N$ of a dyadic subdivision of E, and we may suppose the cubes σ_i to have diameter less than δ. By Lemma 20, there exist automorphisms h and S_2 that carry these cubes into themselves and leave their boundaries fixed such that $hS_2(S_1T)h^{-1}$ is measure-preserving and permutes rigidly in a single cycle the cubes σ_i^* concentric to and with half the diameter of σ_i. But this transformation may be written in the form ST, where $S = hS_2S_1Th^{-1}T^{-1}$. The automorphism S is measure-preserving, since it is a product of two such transformations; it is equal to the identity on $R_1, \cdots, R_K$ and on the boundary of R, since h, S_1, S_2 are; and it has norm less than ϵ, since h, S_1, S_2 have norms less than δ.

THEOREM 12: *Any measure-preserving automorphism T of the r-dimensional unit cube R, $r \geqq 2$, that leaves the boundary fixed can be approximated uniformly by another such automorphism T^* which is locally linear almost everywhere. More precisely, there is a sequence of disjoint cubes interior to R, having total measure one, which under T^* are rigidly permuted among themselves.*

The proof is based on a sequence of modifications of T defined inductively. Let $\delta_1, \delta_2, \cdots$ be a sequence of positive numbers with sum less than a given number $\epsilon > 0$. We shall show that it is possible to find a sequence of automorphisms $S_1, S_2, \cdots$ with the following properties:

1° $S_n \in M_0[R]$, $\rho(S_n, I) < \delta_n$

2° $S_nS_{n-1} \cdots S_1T$ permutes rigidly among themselves a finite number of disjoint cubes $R_1, \cdots, R_{N_n}$ in n cycles, each of these cubes being a sum of interior cubes of a certain dyadic subdivision of R.

3° S_n is equal to the identity on $R_1 + R_2 + \cdots + R_{N_{n-1}}$, $n > 1$.

4° $m(R_1 + \cdots + R_{N_n}) = 1 - \left(1 - \frac{1}{2^r}\right)^n$.

By Lemma 21, there exists $S_1 \in M_0[R]$ with norm less than δ_1 such that S_1T permutes rigidly in a single cycle the cubes $R_1, \cdots, R_{N_1}$ concentric to and with half the diameter of those of a certain dyadic subdivision. These cubes therefore have total measure $1/2^r$. Thus conditions 1° to 4° are satisfied in the case $n = 1$, condition 3° being vacuous. Now suppose that $S_1, \cdots, S_n$ and $R_1, \cdots, R_{N_n}$ have been defined so that conditions 1° to 4° are satisfied for values of $n \leqq k$. Then $T_k = S_kS_{k-1} \cdots S_1T$ fulfills the hypotheses of Lemma 21 and there exists an automorphism $S_{k+1} \in M_0[R]$, with norm less than δ_{k+1}, equal to the identity on $R_1 + \cdots + R_{N_k}$, such that $S_{k+1}T_k$ permutes rigidly in a single cycle the cubes $R_{N_k+1}, \cdots, R_{N_{k+1}}$ concentric to and with half the diameter of those of a certain dyadic subdivision of R that are not contained in $R_1 + \cdots + R_{N_k}$. The total measure of these additional cubes is therefore $\frac{1}{2^r}\left(1 - \frac{1}{2^r}\right)^k$, so that conditions 1° to 4° are now satisfied for values of $n \leqq k + 1$, and the inductive definition of the successive modifications S_n is complete.

From 1° it follows that the limit $S = \lim_{n\to\infty} S_nS_{n-1} \cdots S_1$ exists and has norm less than ϵ. Furthermore, ST is equal to $S_n \cdots S_1T$ on $R_1 + R_2 + \cdots + R_{N_n}$, in virtue of 3°, and so it permutes these cubes rigidly among themselves. Thus under ST the cubes $R_1, R_2, \cdots$ are all rigidly permuted among themselves in finite cycles. Finally, by 4°, these cubes contain almost all points of R. Hence $T^* = ST$ has all the properties described in the theorem.

Bryn Mawr College
The University of Wisconsin

ANNALS OF MATHEMATICS
Vol. 42, No. 4, October, 1941

A REMARK ON NORMAL VARIETIES

BY H. T. MUHLY*

(Received December 22, 1940)

1. In the terminology of the Italian School an algebraic variety is called "normal" if its system of hyperplane sections is complete. O. Zariski applies the term "normal" to an algebraic variety whose associated ring of homogeneous coordinates is integrally closed.[1] The two concepts are not equivalent. Zariski refers to a variety which satisfies the former condition as "normal in the geometric sense" and to one which satisfies the latter condition as "normal in the arithmetic sense."

A variety which is normal in the arithmetic sense is necessarily normal in the geometric sense. Moreover, an r-dimensional variety V_r which is normal in the arithmetic sense has no $(r-1)$-dimensional singularities. A variety which is normal in the geometric sense need not be normal in the arithmetic sense. For example, a plane quartic of genus two is normal in the geometric sense, but has a double point. A curve may be free from singularities and yet not normal in the arithmetic sense. A rational space quartic illustrates this possibility. The validity of all of these assertions is established in Z.

The object of this note is to characterize geometrically those algebraic varieties which are normal in the arithmetic sense. To this end we propose the following theorem: *A necessary and sufficient condition that the r-dimensional algebraic variety V_r be normal in its ambient projective space P_n is that for every integer m the linear system cut out on V_r by the hypersurfaces of order m in P_n be complete.*

In the course of the proof we need the notion of the "character of homogeneity" of an algebraic variety, introduced by Zariski in the paper Z. Let $\xi_0^*, \xi_1^*, \cdots, \xi_n^*$ be the homogeneous coordinates of the general point of a variety V_r in the projective space P_n. The underlying field of constants for P_n is assumed to be algebraically closed and of characteristic zero. We denote this field by K. The integral closure, $K[\zeta_0^*, \zeta_1^*, \cdots, \zeta_h^*]$ of the ring of homogeneous coordinates $K[\xi_0^*, \xi_1^*, \cdots, \xi_n^*]$ in its quotient field Σ^* is denoted by $\bar{o}^*$. Each of the quantities $\zeta_0^*, \zeta_1^*, \cdots, \zeta_h^*$ may be assumed to be homogeneous of positive degree.[2] Zariski proves that there exist integers δ with the following property.

---

* National Research Fellow, 1940–41.

[1] O. Zariski, "Some Results in the Arithmetic Theory of Algebraic Varieties," American Journal of Mathematics, Vol. LXI, No. 2 (April 1939); designated henceforth by Z.

[2] If $\xi_i = \xi_i^*/\xi_0^*$, then $\xi_1, \xi_2, \cdots, \xi_n$ are the non-homogeneous coordinates of the general point of our V_r and $\Sigma(=K(\xi_1, \cdots, \xi_n))$ is the field of rational functions on V_r. The field Σ^* is then equal to $\Sigma(\xi_0)$, a pure transcendental extension of Σ. An element ω^* of Σ^* is said to be homogeneous of degree ν if $\tau \times \omega^* = t^\nu \omega^*$ for every automorphism τ of Σ^* over Σ of the form $\tau : \xi_0^* \to t\xi_0^*$, $t \in K$.

If $\omega_0^*, \omega_1^*, \cdots, \omega_m^*$ denote the possible power products of $\zeta_0^*, \zeta_1^*, \cdots, \zeta_h^*$ which are homogeneous of degree δ and if ω^* is any power product of $\zeta_0^*, \zeta_1^*, \cdots, \zeta_h^*$ of degree of homogeneity $\rho\delta$ (ρ a positive integer) then ω^* is a form of degree ρ in $\omega_0^*, \omega_1^*, \cdots, \omega_m^*$. Any integer δ with this property is called a character of homogeneity of V_r.

2. The necessity of our condition is easily established. The ring of homogeneous coordinates $\mathfrak{o}^* = K[\xi_0^*, \xi_1^*, \cdots, \xi_n^*]$ is integrally closed in its quotient field Σ^* since we assume V_r to be normal in the arithmetic sense. As we have already pointed out, our hypothesis assures us that the linear system $|V|$ which is cut out on V_r by the hyperplanes of P_n is complete. To prove that all of the multiples, $|mV|$, $m = 1, 2, \cdots$, of $|V|$ are also complete we point out that the integer one is a character of homogeneity of V_r since $\mathfrak{o}^*$ is itself integrally closed and since ξ_i^* is homogeneous of degree 1, $i = 0, 1, \cdots, n$. It follows by sections 20 and 21 of Z that if $\omega_0^*, \omega_1^*, \cdots, \omega_h^*$ are the linearly independent power products of degree m (m a positive integer) which occur among all the power products of degree m which can be formed from $\xi_0^*, \xi_1^*, \cdots, \xi_n^*$, then $\omega_0^*, \omega_1^*, \cdots, \omega_h^*$ are the homogeneous coordinates of the general point of a variety V_r' which is normal in the arithmetic sense and which is birationally equivalent to V_r. Moreover, in the birational correspondence between V_r and V_r' the system of sections of V_r with the hypersurfaces of order m is transformed into the system of sections of V_r' with the hyperplanes of its ambient space P_h. Since V_r' is normal in the arithmetic sense the system of hyperplane sections of V_r' is complete. This proves the first half of our theorem.

3. Before proving that our condition is sufficient we list certain well known definitions and results to which we shall have occasion to refer. a) If $\Sigma = K(\xi_1, \cdots, \xi_n)$ is the field of rational functions on V_r, then by a prime divisor of Σ we shall mean an $(r-1)$-dimensional valuation of Σ. If $\mathfrak{J}$ is any finite integral domain in Σ, and if $\mathfrak{P}$ is a prime divisor whose valuation ring $\mathfrak{B}_\mathfrak{P}$ contains $\mathfrak{J}$, then $\mathfrak{P}$ will be said to be of the first kind with respect to $\mathfrak{J}$ if the origin (the prime ideal of elements of positive value in $\mathfrak{J}$) of $\mathfrak{P}$ in $\mathfrak{J}$ is an $(r-1)$-dimensional ideal. If this is not the case then $\mathfrak{P}$ is of the second kind with respect to $\mathfrak{J}$.

b) Let $\xi_0^*, \xi_1^*, \cdots, \xi_n^*$ be the homogeneous coordinates of the general point of a variety V_r in P_n, and let $\mathfrak{o}^* = K[\xi_0^*, \xi_1^*, \cdots, \xi_n^*]$ be the ring associated with $\xi_0^*, \xi_1^*, \cdots, \xi_n^*$. We may assume that $\xi_0^*, \xi_1^*, \cdots, \xi_r^*$ are algebraically independent and that $\xi_{r+1}^*, \cdots, \xi_n^*$ depend integrally on $\xi_0^*, \xi_1^*, \cdots, \xi_r^*$. This situation may always be realized by means of a non-singular projective transformation. If we put $\xi_i = \xi_i^*/\xi_0^*$, $\mathfrak{o}_0 = K[\xi_1, \xi_2, \cdots, \xi_n]$, $\mathfrak{o}_i = K[\xi_1/\xi_i, \cdots, \xi_{i-1}/\xi_i, 1/\xi_i, \xi_{i+1}/\xi_i, \cdots, \xi_n/\xi_i]$, then it is not difficult to see that in each of the rings $\mathfrak{o}_j$, $j = 0, 1, \cdots, r$, the first r coordinates are algebraically independent and that the remaining ones depend integrally on the first r (see section 4). The valuation ring $\mathfrak{B}_\mathfrak{P}$ of any prime divisor, $\mathfrak{P}$, must contain at least one of the rings $\mathfrak{o}_j$, $j = 0, 1, \cdots, r$. Moreover, it is easily seen that if $\mathfrak{o}_\alpha \subset \mathfrak{B}_\mathfrak{P}$ and

$\mathfrak{o}_\beta \subset \mathfrak{B}_\mathfrak{P}$, and if $\mathfrak{P}$ is of the first kind with respect to $\mathfrak{o}_\alpha$ then it is also of the first kind with respect to $\mathfrak{o}_\beta$. Hence we say that $\mathfrak{P}$ is of the first kind (or of the second kind) with respect to the given projective model V_r of Σ.

We consider the set $\mathfrak{S}$ of all prime divisors of Σ which are of the first kind with respect to V_r, and we form the abelian group $\mathfrak{G}(V_r)$ which consists of all finite formal power products $\mathfrak{P}_1^{a_1} \cdot \mathfrak{P}_2^{a_2} \cdots \mathfrak{P}_k^{a_k}$, $\mathfrak{P}_i \in \mathfrak{S}$, α_i an integer, $i = 1, 2, \cdots, k$. The elements of $\mathfrak{G}(V_r)$ are called divisors. The divisor $\mathfrak{A} = \mathfrak{P}_1^{a_1} \cdot \mathfrak{P}_2^{a_2} \cdot \cdots \cdot \mathfrak{P}_k^{a_k}$ is said to be integral if $\alpha_i > 0$, $i = 1, 2, \cdots, k$.

As in the case of functions of one variable, a uniquely determined divisor $\mathfrak{A}(\eta)$ may be associated with each element η in Σ. We have

$$\mathfrak{A}(\eta) = \prod \mathfrak{P}^{v_\mathfrak{P}(\eta)}$$

where the product is extended over all prime divisors in $\mathfrak{S}$. We point out that, since every $\mathfrak{P}$ in $\mathfrak{S}$ is an $(r-1)$-dimensional valuation of Σ, the value group of $\mathfrak{P}$ may be taken to be the group of integers. Those divisors $\mathfrak{A}$ in $\mathfrak{G}(V_r)$ which are associated in this way with elements η in Σ form a subgroup $\mathfrak{H}$ of $\mathfrak{G}(V_r)$. Two divisors $\mathfrak{A}$ and $\mathfrak{B}$ are said to be equivalent if $\mathfrak{A} \cdot \mathfrak{B}^{-1} \in \mathfrak{H}$.

With every prime divisor $\mathfrak{P}$ in $\mathfrak{S}$, one can associate a unique irreducible $(r-1)$-dimensional subvariety $V(\mathfrak{P})$ of V_r. With every integral divisor $\mathfrak{A} = \mathfrak{P}_1^{a_1} \cdot \cdots \cdot \mathfrak{P}_k^{a_k}$, one associates the effective $(r-1)$-dimensional subvariety

$$V(\mathfrak{A}) = \alpha_1 V(\mathfrak{P}_1) + \cdots + \alpha_k V(\mathfrak{P}_k).$$

One then defines the complete system determined by $V(\mathfrak{A})$ to be the totality of all effective curves associated with integral divisors which are equivalent to $\mathfrak{A}$. Zariski's definition of a complete system, which we have used until now, is equivalent to this one in so far as the system of hyperplane sections and its multiples on a given V_r is concerned (see *Z* page 288).

c) LEMMA: *The algebraically independent quantities* $\xi_1, \xi_2, \cdots, \xi_r$ *have divisor representations of the form*

$$\xi_i = \mathfrak{A}_i/\mathfrak{A}, \qquad i = 1, 2, \cdots, r,$$

where $\mathfrak{A}_i$ *and* $\mathfrak{A}$ *are integral divisors and no prime divisor divides both* $\mathfrak{A}_i$ *and* $\mathfrak{A}$. The proof of the lemma runs along the following lines. Let $\mathfrak{P}$ be a prime divisor, of the set $\mathfrak{S}$, which is such that for at least one i the inequality $v_\mathfrak{P}(\xi_i) < 0$ holds. We may assume that $v_\mathfrak{P}(\xi_i) \leqq v_\mathfrak{P}(\xi_j)$, $j = 1, 2, \cdots, r$, for if this is not the case we consider the ξ of least value. The valuation ring $\mathfrak{B}_\mathfrak{P}$ must then contain the ring $\mathfrak{o}_i = K[\xi_1/\xi_i, \cdots, \xi_{i-1}/\xi_i, 1/\xi_i, \xi_{i+1}/\xi_i, \cdots, \xi_n/\xi_i]$. Moreover, since $v_\mathfrak{P}(1/\xi_i) > 0$, $1/\xi_i \equiv 0(\mathfrak{p})$, where $\mathfrak{p}$ is the origin of $\mathfrak{P}$ in $\mathfrak{o}_i$. Since $\mathfrak{p}$ is $(r-1)$-dimensional, $\mathfrak{p}$ cannot divide ξ_j/ξ_i, $j = 1, 2, \cdots, r$. (This follows since $\xi_{r+1}/\xi_i, \cdots, \xi_n/\xi_i$ depend integrally on $\xi_1/\xi_i, \cdots, \xi_{i-1}/\xi_i, 1/\xi_i, \xi_{i+1}/\xi_i, \cdots, \xi_r/\xi_i$.) It therefore follows that $v_\mathfrak{P}(\xi_j/\xi_i) = 0$, or that $v_\mathfrak{P}(\xi_j) = v_\mathfrak{P}(\xi_i)$, $j = 1, 2, \cdots, r$. This proves the lemma.

4. We prove that our condition is sufficient. Let V_r be a variety in the projective space P_n, and let $\xi_0^*, \xi_1^*, \cdots, \xi_n^*$ be the homogeneous coordinates

If $\omega_0^*, \omega_1^*, \cdots, \omega_m^*$ denote the possible power products of $\zeta_0^*, \zeta_1^*, \cdots, \zeta_h^*$ which are homogeneous of degree δ and if ω^* is any power product of $\zeta_0^*, \zeta_1^*, \cdots, \zeta_h^*$ of degree of homogeneity $\rho\delta$ (ρ a positive integer) then ω^* is a form of degree ρ in $\omega_0^*, \omega_1^*, \cdots, \omega_m^*$. Any integer δ with this property is called a character of homogeneity of V_r.

2. The necessity of our condition is easily established. The ring of homogeneous coordinates $\mathfrak{o}^* = K[\xi_0^*, \xi_1^*, \cdots, \xi_n^*]$ is integrally closed in its quotient field Σ^* since we assume V_r to be normal in the arithmetic sense. As we have already pointed out, our hypothesis assures us that the linear system $|V|$ which is cut out on V_r by the hyperplanes of P_n is complete. To prove that all of the multiples, $|mV|$, $m = 1, 2, \cdots$, of $|V|$ are also complete we point out that the integer one is a character of homogeneity of V_r since $\mathfrak{o}^*$ is itself integrally closed and since ξ_i^* is homogeneous of degree 1, $i = 0, 1, \cdots, n$. It follows by sections 20 and 21 of Z that if $\omega_0^*, \omega_1^*, \cdots, \omega_h^*$ are the linearly independent power products of degree m (m a positive integer) which occur among all the power products of degree m which can be formed from $\xi_0^*, \xi_1^*, \cdots, \xi_n^*$, then $\omega_0^*, \omega_1^*, \cdots, \omega_h^*$ are the homogeneous coordinates of the general point of a variety V_r' which is normal in the arithmetic sense and which is birationally equivalent to V_r. Moreover, in the birational correspondence between V_r and V_r' the system of sections of V_r with the hypersurfaces of order m is transformed into the system of sections of V_r' with the hyperplanes of its ambient space P_h. Since V_r' is normal in the arithmetic sense the system of hyperplane sections of V_r' is complete. This proves the first half of our theorem.

3. Before proving that our condition is sufficient we list certain well known definitions and results to which we shall have occasion to refer. a) If $\Sigma = K(\xi_1, \cdots, \xi_n)$ is the field of rational functions on V_r, then by a prime divisor of Σ we shall mean an $(r-1)$-dimensional valuation of Σ. If $\mathfrak{J}$ is any finite integral domain in Σ, and if $\mathfrak{P}$ is a prime divisor whose valuation ring $\mathfrak{B}_\mathfrak{P}$ contains $\mathfrak{J}$, then $\mathfrak{P}$ will be said to be of the first kind with respect to $\mathfrak{J}$ if the origin (the prime ideal of elements of positive value in $\mathfrak{J}$) of $\mathfrak{P}$ in $\mathfrak{J}$ is an $(r-1)$-dimensional ideal. If this is not the case then $\mathfrak{P}$ is of the second kind with respect to $\mathfrak{J}$.

b) Let $\xi_0^*, \xi_1^*, \cdots, \xi_n^*$ be the homogeneous coordinates of the general point of a variety V_r in P_n, and let $\mathfrak{o}^* = K[\xi_0^*, \xi_1^*, \cdots, \xi_n^*]$ be the ring associated with $\xi_0^*, \xi_1^*, \cdots, \xi_n^*$. We may assume that $\xi_0^*, \xi_1^*, \cdots, \xi_r^*$ are algebraically independent and that $\xi_{r+1}^*, \cdots, \xi_n^*$ depend integrally on $\xi_0^*, \xi_1^*, \cdots, \xi_r^*$. This situation may always be realized by means of a non-singular projective transformation. If we put $\xi_i = \xi_i^*/\xi_0^*$, $\mathfrak{o}_0 = K[\xi_1, \xi_2, \cdots, \xi_n]$, $\mathfrak{o}_i = K[\xi_1/\xi_i, \cdots, \xi_{i-1}/\xi_i, 1/\xi_i, \xi_{i+1}/\xi_i, \cdots, \xi_n/\xi_i]$, then it is not difficult to see that in each of the rings $\mathfrak{o}_j$, $j = 0, 1, \cdots, r$, the first r coordinates are algebraically independent and that the remaining ones depend integrally on the first r (see section 4). The valuation ring $\mathfrak{B}_\mathfrak{P}$ of any prime divisor, $\mathfrak{P}$, must contain at least one of the rings $\mathfrak{o}_j$, $j = 0, 1, \cdots, r$. Moreover, it is easily seen that if $\mathfrak{o}_\alpha \subset \mathfrak{B}_\mathfrak{P}$ and

$\mathfrak{o}_\beta \subset \mathfrak{B}_\mathfrak{P}$, and if $\mathfrak{P}$ is of the first kind with respect to $\mathfrak{o}_\alpha$ then it is also of the first kind with respect to $\mathfrak{o}_\beta$. Hence we say that $\mathfrak{P}$ is of the first kind (or of the second kind) with respect to the given projective model V_r of Σ.

We consider the set $\mathfrak{S}$ of all prime divisors of Σ which are of the first kind with respect to V_r, and we form the abelian group $\mathfrak{G}(V_r)$ which consists of all finite formal power products $\mathfrak{P}_1^{a_1} \cdot \mathfrak{P}_2^{a_2} \cdots \mathfrak{P}_k^{a_k}$, $\mathfrak{P}_i \in \mathfrak{S}$, a_i an integer, $i = 1, 2, \cdots, k$. The elements of $\mathfrak{G}(V_r)$ are called divisors. The divisor $\mathfrak{A} = \mathfrak{P}_1^{a_1} \cdot \mathfrak{P}_2^{a_2} \cdot \cdots \cdot \mathfrak{P}_k^{a_k}$ is said to be integral if $a_i > 0$, $i = 1, 2, \cdots, k$.

As in the case of functions of one variable, a uniquely determined divisor $\mathfrak{A}(\eta)$ may be associated with each element η in Σ. We have

$$\mathfrak{A}(\eta) = \prod \mathfrak{P}^{v_\mathfrak{P}(\eta)}$$

where the product is extended over all prime divisors in $\mathfrak{S}$. We point out that, since every $\mathfrak{P}$ in $\mathfrak{S}$ is an $(r-1)$-dimensional valuation of Σ, the value group of $\mathfrak{P}$ may be taken to be the group of integers. Those divisors $\mathfrak{A}$ in $\mathfrak{G}(V_r)$ which are associated in this way with elements η in Σ form a subgroup $\mathfrak{H}$ of $\mathfrak{G}(V_r)$. Two divisors $\mathfrak{A}$ and $\mathfrak{B}$ are said to be equivalent if $\mathfrak{A} \cdot \mathfrak{B}^{-1} \in \mathfrak{H}$.

With every prime divisor $\mathfrak{P}$ in $\mathfrak{S}$, one can associate a unique irreducible $(r-1)$-dimensional subvariety $V(\mathfrak{P})$ of V_r. With every integral divisor $\mathfrak{A} = \mathfrak{P}_1^{a_1} \cdot \cdots \cdot \mathfrak{P}_k^{a_k}$, one associates the effective $(r-1)$-dimensional subvariety

$$V(\mathfrak{A}) = \alpha_1 V(\mathfrak{P}_1) + \cdots + \alpha_k V(\mathfrak{P}_k).$$

One then defines the complete system determined by $V(\mathfrak{A})$ to be the totality of all effective curves associated with integral divisors which are equivalent to $\mathfrak{A}$. Zariski's definition of a complete system, which we have used until now, is equivalent to this one in so far as the system of hyperplane sections and its multiples on a given V_r is concerned (see *Z* page 288).

c) LEMMA: *The algebraically independent quantities $\xi_1, \xi_2, \cdots, \xi_r$ have divisor representations of the form*

$$\xi_i = \mathfrak{A}_i/\mathfrak{A}, \qquad i = 1, 2, \cdots, r,$$

where $\mathfrak{A}_i$ and $\mathfrak{A}$ are integral divisors and no prime divisor divides both $\mathfrak{A}_i$ and $\mathfrak{A}$. The proof of the lemma runs along the following lines. Let $\mathfrak{P}$ be a prime divisor, of the set $\mathfrak{S}$, which is such that for at least one i the inequality $v_\mathfrak{P}(\xi_i) < 0$ holds. We may assume that $v_\mathfrak{P}(\xi_i) \leqq v_\mathfrak{P}(\xi_j)$, $j = 1, 2, \cdots, r$, for if this is not the case we consider the ξ of least value. The valuation ring $\mathfrak{B}_\mathfrak{P}$ must then contain the ring $\mathfrak{o}_i = K[\xi_1/\xi_i, \cdots, \xi_{i-1}/\xi_i, 1/\xi_i, \xi_{i+1}/\xi_i, \cdots, \xi_n/\xi_i]$. Moreover, since $v_\mathfrak{P}(1/\xi_i) > 0$, $1/\xi_i \equiv 0(\mathfrak{p})$, where $\mathfrak{p}$ is the origin of $\mathfrak{P}$ in $\mathfrak{o}_i$. Since $\mathfrak{p}$ is $(r-1)$-dimensional, $\mathfrak{p}$ cannot divide ξ_j/ξ_i, $j = 1, 2, \cdots, r$. (This follows since $\xi_{r+1}/\xi_i, \cdots, \xi_n/\xi_i$ depend integrally on $\xi_1/\xi_i, \cdots, \xi_{i-1}/\xi_i, 1/\xi_i, \xi_{i+1}/\xi_i, \cdots, \xi_r/\xi_i$.) It therefore follows that $v_\mathfrak{P}(\xi_j/\xi_i) = 0$, or that $v_\mathfrak{P}(\xi_j) = v_\mathfrak{P}(\xi_i)$, $j = 1, 2, \cdots, r$. This proves the lemma.

4. We prove that our condition is sufficient. Let V_r be a variety in the projective space P_n, and let $\xi_0^*, \xi_1^*, \cdots, \xi_n^*$ be the homogeneous coordinates

of its general point. We assume that the system $|mV|$ cut out on V_r by the hypersurfaces of order m in P_n is complete, $m = 1, 2, \cdots$. We must prove that the ring $\mathfrak{o}^* = K[\xi_0^*, \xi_1^*, \cdots, \xi_n^*]$ is integrally closed in its quotient field Σ^*. As in section 3, we assume that $\xi_0^*, \xi_1^*, \cdots, \xi_r^*$ are algebraically independent and that $\xi_{r+1}^*, \cdots, \xi_n^*$ depend integrally on $\xi_0^*, \xi_1^*, \cdots, \xi_r^*$. Any element of Σ^* which depends integrally on $\xi_0^*, \xi_1^*, \cdots, \xi_r^*$ is the sum of homogeneous elements which depend integrally on $\xi_0^*, \xi_1^*, \cdots, \xi_r^*$. It is therefore sufficient to prove that $\mathfrak{o}^*$ contains any homogeneous element of Σ^* which depends integrally on $\xi_0^*, \xi_1^*, \cdots, \xi_r^*$.

Let ω^* be an homogeneous element of degree ν with this property and let

$$(1) \qquad \omega^{*h} + a_1(\xi_0^*, \xi_1^*, \cdots, \xi_r^*)\omega^{*h-1} + \cdots + a_h(\xi_0^*, \xi_1^*, \cdots, \xi_r^*) = 0$$

be the equation which expresses the integral dependence of ω^* on $\xi_0^*, \xi_1^*, \cdots, \xi_r^*$. Since every homogeneous integer satisfies an homogeneous equation of integral dependence, we may assume that a_i is a form of degree $i\nu$ in $\xi_0^*, \xi_1^*, \cdots, \xi_r^*$, $i = 1, 2, \cdots, h$ (see Z, section 18). If we divide equation (1) by $\xi_0^{*h\nu}$ we see that the quantity $\omega = \omega^*/\xi_0^{*\nu}$ is an element of $\Sigma\,(= K(\xi_1, \cdots, \xi_n))$ which depends integrally on $\xi_1, \xi_2, \cdots, \xi_r$. On the other hand, division of equation (1) by $\xi_i^{*h\nu}$ shows that ω/ξ_i^{ν} depends integrally on $\xi_1/\xi_i, \cdots, \xi_{i-1}/\xi_i, 1/\xi_i, \xi_{i+1}/\xi_i, \cdots, \xi_r/\xi_i$, $i = 1, 2, \cdots, r$.

Let $\xi_i = \mathfrak{A}_i/\mathfrak{A}$, $i = 1, 2, \cdots, r$, be the divisor decomposition of ξ_i mentioned in 3c. Since the system $|\nu V|$ is complete for every ν, it follows that if $\mathfrak{B}$ is an integral divisor in $\mathfrak{G}(V_r)$ which is equivalent to $\mathfrak{A}^{\nu}$ (in symbols $\mathfrak{B} \sim \mathfrak{A}^{\nu}$), and if $\theta = \mathfrak{B}/\mathfrak{A}^{\nu}$, then $\theta = f(\xi_1, \cdots, \xi_n)$ where f is a polynomial of degree $\leqq \nu$ in the indicated arguments.

If $\theta = \mathfrak{B}/\mathfrak{A}^{\nu}$, then θ is an integral function of $\xi_1, \xi_2, \cdots, \xi_r$, since θ has non-negative value at all prime divisors of $\mathfrak{S}$ for which $\xi_1, \xi_2, \cdots, \xi_r$ have non-negative values. On the other hand, $\theta/\xi_i^{\nu} = \mathfrak{B}/\mathfrak{A}_i^{\nu}$, and consequently θ/ξ_i^{ν} is an integral function of $\xi_1/\xi_i, \cdots, 1/\xi_i, \cdots, \xi_r/\xi_i$. We assert that, conversely, if an integral function θ of $\xi_1, \cdots, \xi_r$ has the property that for some integer ν, θ/ξ_i^{ν} is an integral function of $\xi_1/\xi_i, \cdots, \xi_{i-1}/\xi_i, 1/\xi_i, \xi_{i+1}/\xi_i, \cdots, \xi_r/\xi_i$, $i = 1, 2, \cdots, r$, then θ may be put into the form $\mathfrak{B}/\mathfrak{A}^{\nu}$ where $\mathfrak{B}$ is an integral divisor in $\mathfrak{G}(V_r)$ which is necessarily equivalent to $\mathfrak{A}^{\nu}$. In fact, if $\theta = \mathfrak{B}_0/\mathfrak{A}_0$ is the divisor decomposition of θ, then $\theta/\xi_i^{\nu} = \mathfrak{B}_0\mathfrak{A}^{\nu}/\mathfrak{A}_0\mathfrak{A}_i^{\nu}$. Since θ is an integral function of $\xi_1, \cdots, \xi_r$, no prime divisor of the set $\mathfrak{S}$ can divide $\mathfrak{A}_0$ if it does not divide $\mathfrak{A}$. Moreover, no prime divisor of $\mathfrak{S}$ can occur in $\mathfrak{A}_0$ to a higher power than that to which it occurs in $\mathfrak{A}^{\nu}$, since θ/ξ_i^{ν} is an integral function of $\xi_1/\xi_i, \cdots, \xi_{i-1}/\xi_i, 1/\xi_i, \xi_{i+1}/\xi_i, \cdots, \xi_r/\xi_i$, $i = 1, 2, \cdots, r$. It follows that $\mathfrak{A}^{\nu} = \mathfrak{L} \cdot \mathfrak{A}_0$ where $\mathfrak{L}$ is an integral divisor. We can therefore write $\theta = \mathfrak{B}/\mathfrak{A}^{\nu}$.

Our theorem now follows readily. If ω^* is an element of Σ^* which depends integrally on $\xi_0^*, \xi_1^*, \cdots, \xi_r^*$ and if ω^* is homogeneous of degree ν, then the quantity $\omega = \omega^*/\xi_0^{*\nu}$ is an integral function of $\xi_1, \xi_2, \cdots, \xi_r$ which has the property that ω/ξ_i^{ν} is an integral function of $\xi_1/\xi_i, \cdots, \xi_{i-1}/\xi_i, 1/\xi_i, \xi_{i+1}/\xi_i, \cdots, \xi_r/\xi_i$. The integral function ω therefore admits a divisor representation of the form $\mathfrak{B}/\mathfrak{A}^{\nu}$.

By virtue of the completeness of $|\nu V|$, this implies that $\omega = f(\xi_1, \cdots, \xi_n)$, where f is a polynomial of degree $\leqq \nu$. Hence we see that

$$\omega^* = \xi_0^{*\nu} \cdot f(\xi_1, \cdots, \xi_n) = \varphi_\nu(\xi_0^*, \xi_1^*, \cdots, \xi_n^*)$$

where φ_ν is a form of degree ν. The element ω^* is thus seen to be in $\mathfrak{o}^*$, q.e.d.

As a final remark we point out that there exist curves which are free from singularities and on which the hyperplanes cut out a complete system, but which are nevertheless not normal in the arithmetic sense. In fact in P_3 there exist curves of order seven and of genus four which are free from singularities. On such curves the system of plane sections is complete while the system cut out by the quadric surfaces is not. Thus, even though freedom from singularities and completeness of the system of plane sections are individually necessary for a curve to be normal in the arithmetic sense, they are together not sufficient.

PRINCETON UNIVERSITY AND INSTITUTE FOR ADVANCED STUDY

ANNALS OF MATHEMATICS
Vol. 42, No. 4, October, 1941

ON THE CONNECTION BETWEEN THE ORDINARY AND THE MODULAR CHARACTERS OF GROUPS OF FINITE ORDER*

BY RICHARD BRAUER

(Received October 21, 1940)

Introduction

The representations of groups by matrices with coefficients in a modular field and the corresponding modular group characters have been studied in two earlier papers;[1] the aim of the present paper is to continue this work. Let $\mathfrak{G}$ be a group of finite order g and let p be a rational prime. If $\zeta_1(G)$, $\zeta_2(G)$, $\cdots$ are the (ordinary) characters of $\mathfrak{G}$, and $\varphi_1(G)$, $\varphi_2(G)$, $\cdots$ the modular characters of $\mathfrak{G}$ for p, then we have formulae

$$(*) \qquad \zeta_\mu(G) = \sum_\nu d_{\mu\nu}\varphi_\nu(G)$$

provided that G is a p-regular element of $\mathfrak{G}$, i.e. an element G of an order prime to p. The $d_{\mu\nu}$ are non-negative rational integers, the decomposition numbers of $\mathfrak{G}$ for p. We may say that the group characters ζ_μ of $\mathfrak{G}$ are built up by the modular characters φ_ν, and it is possible to obtain a deeper insight into the nature of the ordinary group characters by the use of the modular characters and their properties. However, it is disturbing that we have to restrict ourselves to p-regular elements. In this paper, we plan to overcome this difficulty. The value $\zeta_\mu(G)$ for elements G of an order divisible by p will be expressed by the modular characters of certain subgroups $\mathfrak{N}_i$ of $\mathfrak{G}$. The corresponding generalized decomposition numbers $d^i_{\mu\nu}$ will not necessarily be rational, but they are integers of a cyclotomic field of an order p^α. The definition of these numbers $d^i_{\mu\nu}$, and the formulae generalizing (*) are given in §1. The numbers $d^i_{\mu\nu}$ can be arranged in the form of a square matrix $\mathbf{D}$ which is non-degenerate, and, apart from the arrangements of the rows and columns, the $d^i_{\mu\nu}$ are invariants of the group $\mathfrak{G}$. The matrix of the group characters of $\mathfrak{G}$ can be written as the product of $\mathbf{D}$ and a matrix A which breaks up completely[2] into the matrices of the modular group characters of $\mathfrak{N}_0$, $\mathfrak{N}_1$, $\mathfrak{N}_2$, $\cdots$, if the rows and columns in all these matrices are suitably arranged. If the modular characters of the groups $\mathfrak{N}_i$ are known, the product of any two columns of $\mathbf{D}$

* Presented to the American Mathematical Society on April 11, 1941.

[1] R. Brauer and C. Nesbitt, On the modular representations of groups of finite order, University of Toronto Studies, Math. Series No. 4, 1937. R. Brauer and C. Nesbitt, On the modular characters of groups. I refer to these two papers by BN 1 and BN 2. The introduction of BN 2 contains a short summary of most of the methods and results of BN 1.

[2] This means that the matrix A contains in its main diagonal the matrices of the modular group characters of $\mathfrak{N}_0$, $\mathfrak{N}_1$, $\cdots$, and zero matrices outside the main diagonal.

can be formed; the Cartan invariants of the groups $\mathfrak{N}_i$ appear as the values of these products (§3). Further, some congruences (mod p) for the decomposition numbers are given which will be of fundamental importance for a later paper.

With every column of **D**, all algebraically conjugate columns appear in **D**. In this manner, the columns of **D** are distributed in "families" of conjugate columns. The number N of such families and the numbers $w_1, w_2, \cdots, w_N$ of members of the individual families are determined in §§6, 7. The numbers obtained are the same as if we distribute the characters ζ_μ into "families" of p-conjugate characters where two characters are said to be *p-conjugate*, if they can be transformed into each other by a change of a primitive $p^{a\text{th}}$ root of unity. Finally, we also obtain these numbers N, w_i, when we distribute the classes of conjugate elements into "families" in a certain manner which can be described in terms of abstract group theory.

NOTATION: $\mathfrak{G}$ denotes a group of finite order g, p is a fixed rational prime number, and we set $g = p^a g'$ with $(p, g') = 1$. A *p-regular* element of $\mathfrak{G}$ is an element whose order is prime to p; the other elements are said to be *p-singular*. Similarly, we denote the classes of conjugate elements as p-regular or p-singular according as the elements of the class are p-regular or p-singular. The normalizor $\mathfrak{N}(G)$ of an element G consists of those elements of $\mathfrak{G}$ which commute with G; the order of $\mathfrak{N}(G)$ will be denoted by $n(G)$. When we speak of a representation or a character of $\mathfrak{G}$ without further attribute, we mean a representation in the field of all complex numbers and the corresponding character. In the case of a representation in a field of characteristic p, we always add the word "modular".

1. Definition of the generalized decomposition numbers

We consider a group $\mathfrak{G}$ of finite order $g = p^a g'$ where p is a prime number and $(g', p) = 1$. Let $P_0 = 1, P_1, P_2, \cdots, P_h$ be a system of elements whose orders are powers of p, such that they all lie in different classes of conjugate elements, but that every element of an order p^α ($\alpha = 0, 1, 2, \cdots$) is conjugate to one of them. Of course, the P_i can be taken from a fixed Sylow subgroup of order p^a. Every p-singular class of $\mathfrak{G}$ contains an element of the form P_iV where i is uniquely determined by the class and where V is a p-regular element of the normalizor $\mathfrak{N}(P_i)$ of P_i. If W is a p-regular element of $\mathfrak{N}(P_j)$ and P_iV and P_jW are conjugate, say

$$G^{-1}P_iVG = P_jW,$$

then by raising this equation to suitable powers, we find

$$G^{-1}P_iG = P_j, \qquad G^{-1}VG = W.$$

Hence $i = j$, G lies in $\mathfrak{N}(P_i)$, and V and W are conjugate in $\mathfrak{N}(P_i)$. Conversely, these conditions imply that P_iV and P_jW are conjugate in $\mathfrak{G}$. In order to obtain a complete system of representatives for the p-singular classes of $\mathfrak{G}$, we have to form P_iV where $i = 0, 1, 2, \cdots, h$, and, for each i, the element V ranges

over a complete system of representatives of the p-regular classes of the group $\mathfrak{N}(P_i)$. Let k_i denote the number of these classes.

Consider a representation $\mathfrak{F}$ of $\mathfrak{G}$. The matrix $\mathfrak{F}(P_i)$ representing a fixed element P_i can be assumed to be of the form

$$\mathfrak{F}(P_i) = \begin{pmatrix} \epsilon_1 I_1 & 0 & \cdots & 0 \\ 0 & \epsilon_2 I_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \epsilon_l I_l \end{pmatrix},$$

where $I_1, I_2, \cdots, I_l$ are unit matrices, and $\epsilon_1, \epsilon_2, \cdots, \epsilon_l$ are distinct $p^{\alpha_{\text{th}}}$ roots of unity; for p^α we may take the order of P_i. The matrix representing an element V_i of $\mathfrak{N}(P_i)$ then breaks up in a corresponding manner,

$$\mathfrak{F}(V_i) = \begin{pmatrix} N^{(1)} & 0 & \cdots & 0 \\ 0 & N^{(2)} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & N^{(l)} \end{pmatrix}.$$

The trace $\operatorname{tr}(\mathfrak{F}(P_iV_i))$ is given by

$$\operatorname{tr}(\mathfrak{F}(P_iV_i)) = \epsilon_1 \operatorname{tr}(N^{(1)}) + \epsilon_2 \operatorname{tr}(N^{(2)}) + \cdots + \epsilon_l \operatorname{tr}(N^{(l)}).$$

If V_i ranges over all elements of $\mathfrak{N}(P_i)$, then $N^{(\lambda)}$, for a fixed λ, will form a representation of $\mathfrak{N}(P_i)$. The trace $\operatorname{tr}(N^{(\lambda)})$ can therefore be expressed as a linear combination of the irreducible characters of $\mathfrak{N}(P_i)$. If V_i is p-regular, these characters again can be expressed by the irreducible modular characters of $\mathfrak{N}(P_i)$.

Let $\varphi_1^i, \varphi_2^i, \cdots$ be the distinct absolutely irreducible modular characters of $\mathfrak{N}(P_i)$ for p; we have k_i of them.[3] Thus we obtain a formula

$$\operatorname{tr}(\mathfrak{F}(P_iV_i)) = \sum d_\nu^i \varphi_\nu^i(V_i)$$

where d_ν^i is an integer of the field of the $p^{\alpha_{\text{th}}}$ roots of unity; it is independent of V_i.

If $\mathfrak{F}_1, \mathfrak{F}_2, \cdots$ are the distinct irreducible representations of $\mathfrak{G}$, and $\zeta_1, \zeta_2, \cdots$ the corresponding characters, we therefore have formulae

$$\zeta_\mu(P_iV_i) = \sum_{\nu=1}^{k_i} d_{\mu\nu}^i \varphi_\nu^i(V_i) \qquad (V_i \text{ in } \mathfrak{N}(P_i), p\text{-regular}). \tag{1}$$

We denote the $d_{\mu\nu}^i$ as the *generalized decomposition numbers* of $\mathfrak{G}$. For $i = 0$, we have $P_0 = 1$, $\mathfrak{N}(P_0) = \mathfrak{G}$, and the $d_{\mu\nu}^0$ are identical with the ordinary decomposition numbers $d_{\mu\nu}$ of $\mathfrak{G}$.[4] In any case, $d_{\mu\nu}^i$ is an integer of the field of the $p^{\alpha_i\text{th}}$ roots of unity where p^{α_i} is the order of P_i.

Let Φ_ν^i be the indecomposable modular character which corresponds to φ_ν^i,

---

[3] BN 1, Theorem III.—BN 2, §8.

[4] BN 1, p. 7.—BN 2, §4.

($\nu = 1, 2, \cdots, k_i$). If $n(P_i)$ is the order of $\mathfrak{N}(P_i)$, then it follows from the orthogonality relations for modular group characters[5] that

$$d^i_{\mu\nu} = \frac{1}{n(P_i)} \sum_{V \text{ in } \mathfrak{N}(P_i)}{}' \zeta_\mu(P_i V)\Phi^i_\nu(V^{-1}), \tag{2}$$

where the dash indicates that V ranges over the p-regular elements of $\mathfrak{N}(P_i)$ only. We arrange these numbers $d^i_{\mu\nu}$ for a fixed i in form of a matrix $D^i = (d^i_{\mu\nu})$ with μ as row index and ν as column index, and set

$$\mathbf{D} = (D^0, D^1, \cdots, D^h). \tag{3}$$

Any column of $\mathbf{D}$ will be denoted as a *d-column of* $\mathfrak{G}$; each of them is given by a pair (i, ν). It follows that the number of such columns is equal to the sum of all k_i, i.e. equal to the full number k of classes of conjugate elements of $\mathfrak{G}$. Hence $\mathbf{D}$ is a square matrix of the same degree k as the matrix Z of the group characters ζ_μ of $\mathfrak{G}$. According to (1), we have a formula $Z = \mathbf{D}A$ where A is a square matrix. Since Z is non-singular, so is $\mathbf{D}$.

THEOREM 1: *The matrix* $\mathbf{D}$ *of the generalized decomposition numbers of* $\mathfrak{G}$ *is non-singular. The matrix of the group characters of* $\mathfrak{G}$ *has the form* $\mathbf{D}A$ *where the matrix* A *breaks up completely*[6] *into the matrices of the modular group characters of* $\mathfrak{N}(P_0), \mathfrak{N}(P_1), \cdots, \mathfrak{N}(P_h)$, *provided that the rows and columns of the matrices are suitably arranged.*

If $\mathbf{D}$ and the modular group characters of all $\mathfrak{N}(P_i)$ are known, then the ordinary group characters ζ_μ of $\mathfrak{G}$ can be found from (1).

2. Change of P_i

For the definition of the generalized decomposition numbers $d^i_{\mu\nu}$, a special system of elements P_i of $\mathfrak{G}$ has been used. We now have to see how the $d^i_{\mu\nu}$ are affected by an admissible change in the choice of P_i, that is when we replace P_i by a conjugate element $G^{-1}P_iG = P_i^*$. Of course, $G^{-1}\mathfrak{N}(P_i)G = \mathfrak{N}(P_i^*)$. If $V \to \mathfrak{F}(V)$, (V in $\mathfrak{N}(P_i)$), is a representation of $\mathfrak{N}(P_i)$, then $V^* \to \mathfrak{F}(GV^*G^{-1})$, V^* in $\mathfrak{N}(P_i^*)$, is a representation of $\mathfrak{N}(P_i^*)$. If $\chi(V)$ is the character of $\mathfrak{F}$, we denote the character of the new representation by χ^G, i.e.

$$\chi^G(V^*) = \chi(GV^*G^{-1}), \qquad V^* \text{ in } \mathfrak{N}(P_i^*). \tag{4}$$

Then

$$(\varphi^i_1)^G, (\varphi^i_2)^G, \cdots, (\varphi^i_{k_i})^G$$

is a complete system of irreducible modular characters of $\mathfrak{N}(P_i^*)$; the corresponding indecomposable characters are

$$(\Phi^i_1)^G, (\Phi^i_2)^G, \cdots, (\Phi^i_{k_i})^G.$$

Since $\zeta_\mu(P_iV) = \zeta_\mu(G^{-1}P_iVG)$, we obtain easily from (2)

[5] BN 1, Theorem IV.—BN 2, (20).
[6] Cf. footnote 2.

THEOREM 2: *If P_i is replaced by $G^{-1}P_iG = P_i^*$, then $d_{\mu\nu}^i = d_{\mu\tau}^{i*}$, where $d_{\mu\tau}^{i*}$ are the decomposition numbers corresponding to this new choice, and where $(\varphi_\tau^i)^G$ is taken as the τ^{th} character of $\mathfrak{N}(P_i^*)$.*

We see that the only possible change is a permutation of the columns of D^i.

THEOREM 3: *The generalized decomposition numbers are invariants of $\mathfrak{G}$.*

3. The orthogonality relations

We form the "unitary product" of two columns of **D**. According to (2), we have

$$(5)\qquad \sum_{\mu=1}^{k} d_{\mu\nu}^i \bar{d}_{\mu\rho}^j = \frac{1}{n(P_i)n(P_j)} \sum_{V \text{ in } \mathfrak{N}(P_i)}' \sum_{W \text{ in } \mathfrak{N}(P_j)}' \Phi_\nu^i(V^{-1})\Phi_\rho^j(W) \sum_\mu \zeta_\mu(P_iV)\overline{\zeta_\mu(P_jW)}.$$

The sum over μ on the right hand side vanishes, if P_iV and P_jW are not conjugate in $\mathfrak{G}$; in the other case its value is $n(P_iV)$. It follows that, for $i \neq j$, the whole expression vanishes. For $i = j$, only the $n(P_i)/n(P_iV)$ elements W are to be taken into account which, in $\mathfrak{N}(P_i)$, are conjugate to V. Hence

$$\sum_{\mu=1}^{k} d_{\mu\nu}^i \bar{d}_{\mu\rho}^i = \frac{1}{n(P_i)} \sum_{V \text{ in } \mathfrak{N}(P_i)}' \Phi_\nu^i(V^{-1})\Phi_\rho^i(V).$$

The right hand side can easily be evaluated.[7] We thus obtain

THEOREM 4: *The generalized decomposition numbers satisfy the equations*

$$(6)\qquad \sum_{\mu=1}^{k} d_{\mu\nu}^i \bar{d}_{\mu\rho}^j = \begin{cases} 0 & \text{for } i \neq j \\ c_{\nu\rho}^i & \text{for } i = j \end{cases}^8$$

where $c_{\nu\rho}^i$ is the Cartan invariant of $\mathfrak{N}(P_i)$ corresponding to the modular characters φ_ν^i, φ_ρ^i.

On multiplying $\mathbf{D}'$ with the conjugate complex matrix $\overline{\mathbf{D}}$ and forming the determinant of the product, we obtain the product of all the determinants $|c_{\kappa\lambda}^i|$, $i = 0, 1, 2, \cdots, h$. All these $|c_{\kappa\lambda}^i|$ are powers of p.[9] With every d-column, the conjugate complex column also appears as a d-column, as is easily seen. Hence $|\overline{\mathbf{D}}| = \pm|\mathbf{D}|$. Since the determinant of the matrix A in theorem 1 is prime to p,[10] we have

THEOREM 5: *The square of the determinant of* **D** *is of the form $\pm p^m$, m a rational integer > 0. The determinant of the second factor A in theorem 1 is prime to p.*

It is not difficult to determine the exact value of the determinant of **D**.

---

[7] BN 1, p. 15.—BN 2, (22).

[8] It should be noted that the $d_{\mu\nu}^i$ are not the ordinary decomposition numbers of $\mathfrak{N}(P_i)$ though they satisfy exactly the same relations.

[9] Cf. R. Brauer, On the Cartan invariants of groups of finite order, Annals of Math. 42, p. 53, 1941.

[10] BN 2, §8.—In BN 1, it is shown that the determinant $|X|$ of the matrix X of the modular group characters of a group is prime to a fixed prime ideal divisor $\mathfrak{p}$ of p, cf. BN 1, (26). The proof given in BN 1, p. 14 for the fact that $|X|$ is prime to p, is not correct. However, this result follows from BN 1, (26), since $|X|^2$ is a rational integer, cf. BN 1, (29).

4. p-conjugate characters

If we replace a primitive g^{th} root of unity ϵ_g by another one ϵ_g^{λ}, $(\lambda, g) = 1$, then every character ζ_μ is transformed into a conjugate character ζ_σ. Choose now $\lambda \equiv 1 \pmod{g'}$ so that the substitution $\epsilon_g \to \epsilon_g^{\lambda}$ amounts to an interchange of the $p^{a_{\text{th}}}$ roots of unity, such that the g'^{th} roots of unity remain unaltered. In this case, we say that the two conjugate characters ζ_μ and ζ_σ are *p-conjugate*. All the characters are distributed into *families of p-conjugate characters.*

If ζ_μ and ζ_σ are p-conjugate, then $\zeta_\mu(G) = \zeta_\sigma(G)$ for p-regular elements G. It follows that ζ_μ and ζ_σ have the same modular constituents (for p). Hence

THEOREM 6: *Two p-conjugate characters ζ_μ and ζ_σ have the same modular constituents; they lie in the same block of characters (for p).*

5. The decomposition numbers corresponding to a block of characters

Let B be a block[11] of characters of $\mathfrak{G}$ (for p). We consider a sum analogous to (6) but where μ ranges only over the values for which ζ_μ belongs to B. We shall say for short that these are the *indices in* B. Similar to (5) we have

$$\sum_{\mu \text{ in B}} d_{\mu\nu}^{i} \bar{d}_{\mu\rho}^{j} = \sum_{V}{}' \sum_{W}{}' \frac{1}{n(P_i)} \frac{1}{n(P_j)} \Phi_\nu^i(V^{-1}) \Phi_\rho^j(W) \sum_{\mu \text{ in B}} \zeta_\mu(P_i V) \overline{\zeta_\mu(P_j W)}.$$

As the $d_{\mu\nu}^{i}$, the whole sum is an integer of the field of the $p^{a_{\text{th}}}$ roots of unity. But with every ζ_μ all its p-conjugates appear, and the expression on the right hand side shows that the sum is a rational integer.

Now collect the terms for which V lies in a fixed class of $\mathfrak{N}(P_i)$ and W in a fixed class of $\mathfrak{N}(P_j)$. Since these classes consist of $n(P_i)/n(P_iV)$ and of $n(P_j)/n(P_jW)$ elements respectively, and all the corresponding terms have the same value, we find

$$\sum_{\nu} d_{\mu\nu}^{i} \bar{d}_{\mu\rho}^{j} = \sum_{V}{}'' \sum_{W}{}'' [\Phi_\nu^i(V^{-1})/n(P_i V)][\Phi_\rho^j(W)/n(P_j W)] \sum_{\mu \text{ in B}} \zeta_\mu(P_i V) \overline{\zeta_\mu(P_j W)} \tag{7}$$

where V and W range over certain elements of $\mathfrak{N}(P_i)$ and $\mathfrak{N}(P_j)$ respectively. The numbers in the square brackets are p-integers.[12] If $\mathfrak{p}$ is a prime ideal divisor of p in the field generated by the characters ζ_μ, then

$$\zeta_\mu(P_i V) \equiv \zeta_\mu(V) \pmod{\mathfrak{p}}.$$

On the other hand, if $j > 0$ and, therefore, $P_j W$ p-singular, we have

$$\sum_{\mu \text{ in B}} \zeta_\mu(V) \overline{\zeta_\mu(P_j W)} = 0.^{13}$$

---

[11] BN 1, §§6, 7.—BN 2, §9.

[12] BN 1, theorem V.—This theorem is a consequence of BN 2, (16) and (17).

[13] BN 1, theorem VIII.—This can also be seen from BN 2, (28), and the formulae (1) and (6) of the present paper.

Hence for $j > 0$ the sum (7) is divisible by $\mathfrak{p}$, and since it is rational, it is divisible by p.

THEOREM 7: *If $P_j \neq 1$, i.e. if $j > 0$, then for every block* B

$$\sum_{\mu \text{ in } B} d^i_{\mu\nu} \bar{d}^i_{\mu\rho} \equiv 0 \pmod{\mathfrak{p}}; \tag{8}$$

the left side is a rational integer.

Assume that the blocks $B = B_\lambda$ consists of the ordinary characters $\zeta_1, \zeta_2, \cdots, \zeta_x$ and the modular characters $\varphi_1, \varphi_2, \cdots, \varphi_y$, $(\varphi_\nu = \varphi^0_\nu)$ and form the matrix $D_\lambda = (d_{\rho\sigma}) = (d^0_{\rho\sigma})$, $(\rho = 1, 2, \cdots, x; \sigma = 1, 2, \cdots, y)$. The matrix D breaks up completely into $D_1, D_2, \cdots$ corresponding to the different blocks.[14] From (6), it follows that

$$\sum_{\mu \text{ in } B} d^i_{\mu\nu} d^0_{\mu\rho} = \begin{cases} c^0_{\nu\rho} = c_{\nu\rho}, & \text{if } i = 0, \text{ and } \varphi_\nu, \varphi_\rho \text{ both belong to B.} \\ 0, & \text{in all other cases.} \end{cases} \tag{8*}$$

This supplements (8).

If $\zeta_1, \zeta_2, \cdots, \zeta_x$ belong to w different families of p-conjugate characters, then we arrange the ζ_μ so that $\zeta_1, \zeta_2, \cdots, \zeta_w$ all belong to different families. Assume that the family of ζ_ν consists of r_ν characters and set

$$\tilde{D}_\lambda = (d_{\rho\sigma}) \qquad (\rho = 1, 2, \cdots, w; \sigma = 1, 2, \cdots, y). \tag{9}$$

If ζ_μ and ζ_ρ are p-conjugate, then $d_{\mu\sigma} = d_{\rho\sigma}$ by theorem 6. Therefore, D_λ has the same rows as $\tilde{D}_\lambda$, but the μ^{th} row of $\tilde{D}_\lambda$ appears r_μ times in D_λ. From (8*) and (1), it follows that

$$\sum_{\mu=1}^{x} d_{\mu\nu} \zeta_\mu(S) = 0 \qquad \text{(for all } p\text{-singular } S \text{ of } \mathfrak{G}).$$

Here, p-conjugate characters have the same coefficients. If we denote by $\tilde{\zeta}_\mu$ the sum of all characters which are p-conjugate to ζ_μ (including ζ_μ), the equation can be written in the form

$$\sum_{\mu=1}^{w} d_{\mu\nu} \tilde{\zeta}_\mu(S) = 0 \qquad \text{(for all } p\text{-singular } S \text{ of } \mathfrak{G}). \tag{10}$$

We set $d_\mu = \sum_\nu d_{\mu\nu}\omega_\nu$ for any fixed numbers $\omega_1, \cdots, \omega_y$. Then

$$\sum_{\mu=1}^{w} d_\mu \tilde{\zeta}_\mu(S) = 0 \qquad \text{(for all } p\text{-singular } S \text{ of } \mathfrak{G}). \tag{10a}$$

Every character $\zeta_\mu(G)$ of $\mathfrak{G}$ may be considered as a character of the p-Sylow-subgroup $\mathfrak{P}$ of $\mathfrak{G}$, when we take G as an element P of $\mathfrak{P}$. If, in this sense, ζ_μ contains the 1-character [1] of $\mathfrak{P}$ q_μ times, the same will hold for the p-conjugate characters and, therefore, $\tilde{\zeta}_\mu$ will contain [1] exactly $r_\mu q_\mu$ times. The expression $\chi(P) = \sum d_\mu \tilde{\zeta}_\mu(P)$ is a linear combination of the characters of $\mathfrak{P}$, and [1] appears

[14] BN 1, theorem VII.—BN 2, (28).

in it with the coefficient $\sum d_\mu r_\mu q_\mu$. But, from (10a), $\chi(P) = 0$ for $P \neq 1$. If z_μ is the degree of ζ_μ, then $\chi(1) = \sum r_\mu d_\mu z_\mu$. The orthogonality relations for ordinary group characters give $(1/p^a) \sum r_\mu d_\mu z_\mu$ as the coefficient of [1] in χ, and hence

$$p^a \sum_{\mu=1}^{w} r_\mu d_\mu q_\mu = \sum_{\mu=1}^{w} r_\mu d_\mu z_\mu. \tag{11}$$

If only one of the numbers d_μ is different from 0, (11) becomes $p^a q_\mu = z_\mu$, and then the character ζ_μ is of the highest kind.[15] If we exclude this case, it follows that it is impossible to determine $\omega_1, \omega_2, \cdots, \omega_y$ so that only one of the d_μ does not vanish. This implies that the rank of $\tilde{D}_\lambda$ is smaller than w. But $\tilde{D}_\lambda$ has the same rank as D_λ, and this rank is y.[16] Hence

THEOREM 8: *Every block* B *which is not of the highest kind contains more families of ordinary characters than it contains modular characters*: $w > y$.[17]

Each relation (10) must contain at least two $\tilde{\zeta}_\mu$. This gives

THEOREM 9: *If the block* B *is not of the highest kind, then each of its modular constituents appears in at least two characters* ζ_μ *of* B *which are not p-conjugate.*

From (10) it also follows that

$$\sum_{\mu=1}^{w} \zeta_\mu(R)\tilde{\zeta}_\mu(S) = 0 \tag{12}$$

for any p-regular R and any p-regular S.

The blocks of highest kind consist of exactly one ζ_μ whose degree is divisible by p^a, and each such ζ_μ forms a block of highest kind.[18] Since such a ζ_μ vanishes for all p-singular elements, (2) gives

THEOREM 10: *If* ζ_μ *forms a block of highest kind, then* $d^i_{\mu\nu} = 0$ *for all* $i > 0$ *and all* ν.

6. The permutation lemma

We now derive a simple lemma which we shall need. Consider a matrix $M = (m_{ij})$ with u rows and v columns. Every permutation A of the rows of M can be effected by left-multiplication of M with a suitable "permutation matrix" P_A of degree u which in every row and in every column has one coefficient 1 and $m - 1$ coefficients 0. Similarly, every permutation B of the columns of M can be effected by right-multiplication of M with a suitable permutation matrix Q_B of degree v. We prove

LEMMA 1: *Let* M *be a non-singular matrix. If there exists a permutation* A *of the rows of* M *and a permutation* B *of the columns of* M *such that both carry* M *into the same matrix, then the cycles of the permutation* A *have the same lengths as those of* B. *In particular,* A *and* B *have the same number of cycles.*

---

[15] BN 2, theorem 1.

[16] BN 1, p. 21.—BN 2, (29), (15) and (14).

[17] This improves the inequality $x > y$ given in BN 2.

[18] Cf. footnote 15.

PROOF: According to the assumption, we have

$$P_A M = M Q_B . \tag{13}$$

Since M is non-singular, P_A and Q_B have the same characteristic roots. To each cycle of length r of A, there correspond the r r^{th} roots of unity as characteristic roots of P_A.[19] On comparing the roots of P_A and Q_B, starting with the maximal r, we readily obtain lemma 1. Similarly, we prove

LEMMA 2: *Let M be a rectangular matrix whose columns are linearly independent, and assume that there exists a permutation A of the rows and a permutation B of the columns of M which both carry M into the same matrix. If B has a cycle of length r then A has a cycle whose length is divisible by r.*

PROOF: Again, (13) holds. Schur's lemma shows that Q_B is a constituent of P_A so that the characteristic roots of Q_B appear among those of P_A. On comparing these roots, we obtain lemma 2. We also have the result

LEMMA 3: *Let M be a non-singular matrix of degree m, and let $\mathfrak{A}$ and $\mathfrak{B}$ be two permutation groups of degree m which are both homomorphic to the same group $\mathfrak{T}$. If for every T in $\mathfrak{T}$, the corresponding element A_T of $\mathfrak{A}$, applied to the rows of M, and the corresponding element B_T of $\mathfrak{B}$, applied to the columns of M, both carry M into the same matrix, then the number of systems of transitivity is the same for $\mathfrak{A}$ and for $\mathfrak{B}$.*

PROOF: Again (13) will hold for corresponding $A = A_T$ and $B = B_T$. All we have to show is that the number $\tau_{\mathfrak{A}}$ of systems of transitivity is an invariant, if $\mathfrak{A}$ is interpreted as a group of linear transformations, and similarity transformations of $\mathfrak{A}$ are performed. But this follows from the fact that $\tau_{\mathfrak{A}}$ is the number of 1-constituents of the linear group $\mathfrak{A}$.

7. Families of characters, classes, and d-columns

The results of §6 can be applied when M is the matrix of group characters of $\mathfrak{G}$. We may construct corresponding permutations A and B in the following manner. Let ϵ_g be a primitive g^{th} root of unity, and let λ be a rational integer which is prime to g. The substitution $T_\lambda : \epsilon_g \to \epsilon_g^\lambda$ carries every character χ of $\mathfrak{G}$ into a conjugate character $\chi^{(\lambda)}$, we have

$$\chi^{(\lambda)}(G) = \chi(G^\lambda). \tag{14}$$

On the other hand, the substitution $G \to G^\lambda$ carries every class of conjugate elements $\mathfrak{C}$ into a new class $\mathfrak{C}^{(\lambda)}$. Then (14) shows that the value of χ for $\mathfrak{C}^{(\lambda)}$ is the same as the value of $\chi^{(\lambda)}$ for $\mathfrak{C}$. Hence the permutation $A: \chi \to \chi^{(\lambda)}$ of the rows of M, and the permutation $B: \mathfrak{C} \to \mathfrak{C}^{(\lambda)}$ of the columns both carry M into the same matrix.

We are interested in the case that $\lambda \equiv 1 \pmod{g'}$. Then χ and $\chi^{(\lambda)}$ belong to

---

[19] A modification is necessary, if the underlying field is modular, but the lemma remains valid. The same is true for lemma 2 and lemma 3. We shall use the lemmas only in the case of a non-modular field.

the same family of p-conjugate classes (§4). We shall also say that the classes $\mathfrak{C}$ and $\mathfrak{C}^{(\lambda)}$ belong to the same family of classes ($\lambda \equiv 1 \pmod{g'}$). If $\mathfrak{C}$ contains the element P_iV (V in $\mathfrak{N}(P_i)$, p-regular), then $\mathfrak{C}^{(\lambda)}$ contains $P_i^{\lambda}V$ in this case.

Before formulating the results, we also consider the matrix $\mathbf{D}$, (3). With every column $d_{\mu\nu}^{i}$, all the algebraically conjugate columns will appear. Indeed, the substitution T_λ, ($\lambda \equiv 1 \pmod{g'}$) results in the replacing of P_i in (2) by P_i^{λ}, and on account of Theorem 2, this new column can again be expressed in the form $d_{\mu\tau}^{j}$. The d-columns thus appear distributed into families of algebraically conjugate d-columns. The effect of the substitution T_λ on $\mathbf{D}$ then consists of a permutation B^* of the columns; the members of each family are interchanged among themselves.

On the other hand, the effect of T_λ on $\mathbf{D}$ can also be described by the permutation $A: \chi \to \chi^{(\lambda)}$ of the rows of $\mathbf{D}$ as follows from (2). Hence the assumptions of lemma 1 are also satisfied for $M = \mathbf{D}$ and the permutations A and B^*.

Let $\mathfrak{T}$ be the group of all substitutions T_λ with $\lambda \equiv 1 \pmod{g'}$. We then have homomorphic groups $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{B}^*$ consisting of the A, the B, and the B^* respectively. In each of the three cases, a system of transitivity corresponds exactly to a family (of characters, classes, or d-columns). Hence lemma 3 gives

THEOREM 11: *The number of distinct families is the same for each of the three kinds of families: Families of p-conjugate characters, families of classes, and families of conjugate d-columns.*[20]

Next let p be an odd prime. Thus $\mathfrak{T}$ is cyclic, and a primitive element is obtained by taking for λ a primitive root (mod p^a), ($\lambda \equiv 1 \pmod{g'}$). For this T_λ, the lengths of the cycles of the permutation A are the numbers of members belonging to the different families of characters. A similar statement holds for B and B^*. Lemma 1 now yields

THEOREM 12: *Let $p \neq 2$. If the different families of characters contain $r_1, r_2, \cdots, r_f$ members respectively, if the different families of classes contain $s_1, s_2, \cdots, s_f$ respectively, and if the different families of d-columns contain $t_1, t_2, \cdots, t_f$ members respectively, then the three sets $(r_1, r_2, \cdots, r_f)$, $(s_1, s_2, \cdots, s_f)$, $(t_1, t_2, \cdots, t_f)$ are identical apart from the arrangement.*[21]

Remark: The k_0 p-regular classes form each a family of its own, $s_\kappa = 1$. Similarly, the k_0 d-columns of $D^0 = D$ form each a family of its own, $t_\kappa = 1$.

There is another case in which lemma 1 and lemma 3 can be applied. Every automorphism of $\mathfrak{G}$ permutes the characters, the classes, and the d-columns, and again the assumptions of the lemmas are satisfied. It seems unnecessary to formulate the results explicitly.

UNIVERSITY OF TORONTO,
TORONTO, CANADA.

[20] For the first two kinds of families, compare the similar statement and proof in W. Burnside, Theory of groups of finite order, 2nd. ed., Cambridge 1911, p. 315, theorem VI.

[21] For $p = 2$ this will hold, if G does not contain elements of order 8.

ANNALS OF MATHEMATICS
Vol. 42, No. 4, October, 1941

INVESTIGATIONS ON GROUP CHARACTERS*

BY RICHARD BRAUER

(Received November 19, 1940)

Introduction

Let $\mathfrak{G}$ be a finite group of order g. It is well known that the distinct irreducible representations $\mathfrak{Z}_1, \mathfrak{Z}_2, \cdots, \mathfrak{Z}_k$ of $\mathfrak{G}$ in the field of complex numbers can be so chosen that their coefficients belong to an algebraic number field Ω of finite degree. Furthermore, if p is a fixed rational prime number, we may assume that the coefficients are p-integers, i.e. are of the form α/β where α and β are integers of Ω and β is prime to p. Let $\mathfrak{P}$ be a fixed prime ideal divisor of p in Ω. If every coefficient of $\mathfrak{Z}_\kappa$ is replaced by its residue class (mod $\mathfrak{P}$), then we obtain a modular representation $\overline{\mathfrak{Z}}_\kappa$ with coefficients in a field of characteristic p.

It was shown by Dickson and Speiser[1] that the ordinary theory of group characters remains valid for modular group characters, if the prime p does not divide the group order g. Every modular representation then is completely reducible; all the distinct, absolutely irreducible modular representations are given by $\overline{\mathfrak{Z}}_1, \overline{\mathfrak{Z}}_2, \cdots, \overline{\mathfrak{Z}}_k$.

In a recent paper,[2] C. Nesbitt and the author obtained results which may be considered as generalizations of these theorems. Let p be any rational prime, and assume that p^a is the highest power of p which divides g, say

$$g = p^a g', \qquad (p, g') = 1.$$

We then considered representations $\mathfrak{Z}_\kappa$ whose degree is divisible by p^a. It was shown that $\overline{\mathfrak{Z}}_\kappa$ is absolutely irreducible as a modular representation. Whenever $\overline{\mathfrak{Z}}_\kappa$ appears as a constituent of a modular representation $\mathfrak{F}$, then $\mathfrak{F}$ breaks up completely into $\overline{\mathfrak{Z}}_\kappa$ and another constituent $\mathfrak{A}$ (reducible or irreducible)

$$\mathfrak{F} \sim \begin{pmatrix} \overline{\mathfrak{Z}}_\kappa & 0 \\ 0 & \mathfrak{A} \end{pmatrix}.$$

None of the representations $\overline{\mathfrak{Z}}_\lambda$ for $\lambda \neq \kappa$ contains $\overline{\mathfrak{Z}}_\kappa$ as a constituent. Since every irreducible modular representation appears as a constituent of at least one of the representations $\overline{\mathfrak{Z}}_1, \overline{\mathfrak{Z}}_2, \cdots, \overline{\mathfrak{Z}}_k$, this, (in the case $a = 0$, i.e. $g \not\equiv 0 \pmod{p}$), actually yields the theorem of Dickson and Speiser.

---

* Presented to the American Mathematical Society on September 5, 1941.

[1] L. E. Dickson, Trans. Am. Math. Soc. **3**, p. 285, 1902. A. Speiser, *Theorie der Gruppen von endlicher Ordnung* 3rd ed. Berlin 1937, §71.

[2] *On the modular characters of groups*, Ann. of Math., **42**, p. 556, 1941. I refer to this paper as BN.

In this paper, I study representations $\mathfrak{Z}_\kappa$ whose degree z_κ is divisible by p^{a-1} but not by p^a. If the order g of $\mathfrak{G}$ is divisible by p to the first power only, then every representation $\mathfrak{Z}_\kappa$ is either of this type, or of the "highest" type, $z_\kappa \equiv 0 \pmod{p^a}$, which was studied in the former paper. Our results will enable us to derive a great number of properties of the characters of groups $\mathfrak{G}$ of such an order $g = pg'$; these properties form a powerful weapon in the investigation of these groups.[3]

Our first result concerning representations $\mathfrak{Z}_\kappa$ of a degree $z_\kappa \equiv 0 \pmod{p^{a-1}}$ is that the degree of all those representations $\mathfrak{Z}_\lambda$ which belong to the same block[4] B as $\mathfrak{Z}_\kappa$, is also divisible by p^{a-1}. I mention here the following application of this theorem: The only simple group of an order $4p^aq^b$, (p, q primes) with $a \leqq 2$ is the simple group of order 60; the only simple groups of an order $3p^aq^b$ (p, q primes) with $a \leqq 2$ are the simple groups of order 60 and 168. (§9.)

With the block B containing a representation $\mathfrak{Z}_k$ of degree $z_\kappa \equiv 0 \pmod{p^{a-1}}$, $z_\kappa \not\equiv 0 \pmod{p^a}$, we associate a linear graph. Every vertex V_λ corresponds to a family of p-conjugate characters[5] ζ_λ, ζ_λ', $\cdots$ of B, every edge S_μ to a modular character φ_μ of B, and the edge S_μ contains V_λ, if φ_μ is a modular constituent of ζ_λ (and so of all its p-conjugates). It will be shown that every S_λ contains only two vertices, further that the complex actually is a tree T. Since we also have the result that φ_μ never appears with higher multiplicity than 1 in an ordinary irreducible character, the tree describes the complete structure of the block B, if at every vertex V_λ the number r_λ of characters in the family of ζ_λ is indicated. If T and these numbers r_λ are given, the decomposition numbers and hence the Cartan invariants corresponding to the block can be easily obtained. Of course, there exist two vertices of T which lie on only one edge. Consequently, the block B contains at least two characters which are not p-conjugate, and which remain irreducible, when considered as modular characters.

If $\mathfrak{Z}_\lambda$ is any irreducible representation of $\mathfrak{G}$, if z_λ is its degree and r_λ the number of its p-conjugates, then we can show that $r_\lambda z_\lambda \not\equiv 0 \pmod{p^{a+1}}$.[6] If $r_\lambda z_\lambda \equiv 0 \pmod{p^a}$, then $z_\lambda \equiv 0 \pmod{p^a}$, i.e. $\mathfrak{Z}_\lambda$ is of the highest kind.

In the case $z_\lambda \equiv 0 \pmod{p^{a-1}}$, $z_\lambda \not\equiv 0 \pmod{p^a}$, it follows that r_λ divides $p - 1$. Besides, we have the relation

$$r_1 r_2 \cdots r_w \left(\frac{1}{r_1} + \frac{1}{r_2} + \cdots + \frac{1}{r_w}\right) = p$$

for the numbers r_λ belonging to the different vertices V_λ of the tree T.

In the last three sections, the arrangement of the modular constituents of an

---

[3] Cf. R. Brauer, *On groups whose order contains a prime factor to the first order*, to appear soon.

[4] BN, §9. Cf. also R. Brauer and C. Nesbitt, *On the modular representations of groups of finite order*, University of Toronto Studies, Math. Series No. 4, 1937.

[5] For the definition of p-conjugate characters, cf. the list of notations at the end of this introduction.

[6] This improves the theorem that the degree of an irreducible representation divides the order of the group.

irreducible representation $\mathfrak{Z}_\lambda$ of a degree $z_\lambda \equiv 0 \pmod{p^{a-1}}$ is discussed. We assume here that the splitting field Ω is normal over the field of rational numbers and that the order of ramification e of p in Ω is equal to the number r_λ of p-conjugates; there exist fields Ω satisfying these conditions, but e can never be smaller than r_λ. It turns out that the arrangement of the modular constituents of $\mathfrak{Z}_\lambda$ is uniquely determined apart from a cyclic permutation of the constituents. If we restrict ourselves to the application of similarity transformations M whose coefficients are $\mathfrak{P}$-integers and whose determinant is prime to $\mathfrak{P}$, then the cyclic permutation must be the identity. The class of all representations $\mathfrak{Z}_\nu$ with $\mathfrak{P}$-integral coefficients in Ω, which are similar to $\mathfrak{Z}_\lambda$, splits into j subclasses of representations which are similar in this narrower sense. Here, j is the number of modular constituents of $\mathfrak{Z}_\lambda$, and each of the subclasses corresponds to one of the j cyclic permutations of the modular constituents.

NOTATION: $\mathfrak{G}$ is a group of order $g = p^a g'$ where p is a fixed prime and $(g', p) = 1$. The words "representation" and "character" always refer to representations in the field of all complex numbers and their characters, unless the word "modular" is added, in which case we mean a representation in a field of characteristic p. The distinct irreducible representations of $\mathfrak{G}$ are denoted by $\mathfrak{Z}_1, \mathfrak{Z}_2, \cdots, \mathfrak{Z}_k$, their characters by $\zeta_1, \zeta_2, \cdots, \zeta_k$ and their degrees by $z_1, z_2, \cdots, z_k$. Two characters ζ_κ and ζ_λ are p-conjugate if ζ_κ can be carried into ζ_λ by a change of the primitive p^ath roots of unity which leaves the g'th roots of unity unaltered. Then $\zeta_1, \zeta_2, \cdots, \zeta_k$ are distributed into "families" of p-conjugate characters; the number of members of the family of ζ_λ is usually denoted by r_λ. If G is an element of $\mathfrak{G}$, then $\zeta_\lambda(G)$ is the value of ζ_λ for this element G. The distinct, absolutely irreducible modular characters of G are denoted by $\varphi_1, \varphi_2, \cdots$. An element G of $\mathfrak{G}$ is p-regular if its order is prime to p. If its order is divisible by p, then G is p-singular.

1. Construction of a suitable splitting field[7]

Let $\mathfrak{G}$ be a group of finite order g, and consider an irreducible representation $\mathfrak{Z}$ of the group $\mathfrak{G}$ in the field of all complex numbers. We want to construct a splitting field Λ for $\mathfrak{Z}$ such that Λ is normal over the field P of rational numbers, and that the order of ramification e of a given prime p in Λ is as small as possible. Of course, Λ must contain the field $\mathrm{Z} = \mathrm{P}(\zeta)$ generated by the character ζ of $\mathfrak{Z}$, and this fact imposes a condition on e. The following lemma shows that, for suitable Λ, the number e has the smallest value which is compatible with this condition.

LEMMA 1: *There exist splitting fields of the form* $\mathrm{Z}(\tau)$ *where* τ *is a root of unity,* $\tau^s = 1$, *of an order* s *which is prime to* p.

PROOF: Let $\mathfrak{q}$ be a prime ideal of Z and $\mathfrak{Q}$ be a corresponding prime ideal of

[7] For the concepts of the theory of algebras used in §1, cf. M. Deuring, *Algebren*, Berlin 1935, and A. A. Albert, *Structure of algebras*, New York 1939.

$Z(\tau)$. According to a theorem of Hasse,[8] the field $Z(\tau)$ will be a splitting field, if for every $\mathfrak{q}$ the $\mathfrak{q}$-index $m_\mathfrak{q}$ of Z divides the $\mathfrak{Q}$-degree $n_\mathfrak{q}$ of $Z(\tau)$. Let $\mathfrak{q}$ be one of the prime ideals for which $m_\mathfrak{q} \neq 1$, and let r^ν be the highest power of the rational prime r which divides $m_\mathfrak{q}$. We shall show below that we can find a root of unity $\vartheta = \vartheta(\mathfrak{q}, r^\nu)$ of an order $j = j(\mathfrak{q}, r^\nu)$ prime to p, such that for every prime divisor $\mathfrak{Q}_\vartheta$ of q in $Z(\vartheta)$, the $\mathfrak{Q}$-degree is divisible by r^ν. If we then adjoin the numbers $\vartheta(\mathfrak{q}, r^\nu)$ for all discriminant divisors $\mathfrak{q}$, and all prime powers r^ν dividing $m_\mathfrak{q}$, a field $Z(\tau)$ with the desired property is obtained.

The construction of ϑ is immediate for infinite prime ideals; any imaginary root of unity of an order prime to p can be taken. Let $\mathfrak{q}$ be finite and let q be the rational prime divisible by $\mathfrak{q}$. It will be sufficient to find a positive rational integer j which is prime to p, such that in a field of the j^{th} roots of unity either the order of ramification e_q or the degree f_q of the prime ideal divisors of q is divisible by a preassigned prime power r^t.[9] We distinguish several cases:

(a) if $r \neq p$, $r \neq q$ then assume that q belongs to the exponent ρ (mod r). We set $j = r^\lambda$ where λ is a positive rational integer, and have $q^{\rho j} \equiv 1 \pmod{j}$. The degree f_q equals the exponent to which q belongs (mod j). Hence $\rho \mid f_q$, but $f_q \mid \rho j$ and consequently, f_q is of the form ρr^σ. For sufficiently large λ, we shall have $\sigma \geqq t$, i.e. $r^t \mid f_q$.

(b) If $r = q$, $q \neq p$, take $j = r^{t+1}$. Then $e_q = r^t(r-1)$.

(c) If $r = p$, $q \not\equiv 1 \pmod{p}$, take $j = q^{r^t} - 1$. Then $(j, p) = 1, f_q = r^t$.

(d) If $r = p$, $q \equiv 1 \pmod{p}$, we may write $q^{r^\lambda} - 1 = c_\lambda r^{b_\lambda}$ where $(c_\lambda, r) = 1$, $b_\lambda > 0$.[10] Raising this equation to the r^{th} power we easily obtain $c_{\lambda-1} < c_\lambda$. For $j = c_\lambda$, $\lambda \geqq t$, we have $f_q = r^\lambda \geqq r^t$. In all these cases j is prime to p. This finishes the proof of the lemma.

Without restriction, we may assume that all g^{th} roots of unity belong to $Z(\tau)$. We may further assume that τ is so chosen that it can be used simultaneously for all irreducible representations of $\mathfrak{G}$. If we set $K = P(\tau)$, we have

LEMMA 1*: *There exists a field* K *over the field of rational numbers with the following properties*:

(α) *The field* K *contains the* g^{th} *roots of unity.*

(β) *The prime* p *is not ramified in* K: $(p) = \mathfrak{p}\mathfrak{m}$, *where* $\mathfrak{p}$ *is a prime ideal of* K, *and* $(\mathfrak{p}, \mathfrak{m}) = (1)$.

(γ) *Every irreducible representation* $\mathfrak{Z}$ *of* $\mathfrak{G}$ *can be written in the field* $K(\zeta)$ *obtained from* K *by adjoining the character* ζ *of* $\mathfrak{Z}$.

Two characters ζ and ζ' are algebraically conjugate with regard to K, if and only if they are p-conjugate. The degree r of $K(\zeta)$ with regard to K is, therefore equal to the number of characters in the family of ζ.

Let ϵ be a p^αth root of unity such that ζ lies in $K(\epsilon) = \Omega$; we may always take

---

[8] Cf. the books mentioned in footnote 7 or the original paper of Hasse, Math. Ann. **107**, p. 731 (1933).

[9] Since the field Z is fixed, the field $Z(\vartheta)$ will satisfy the condition above, if t is taken large enough.

[10] If $r = 2$, choose $\lambda \geqq 2$.

$1 \leqq \alpha \leqq a$. Let $\mathfrak{P}$ be the prime ideal divisor of $\mathfrak{p}$ in Ω.[11] The representation $\mathfrak{Z}$ can be written with $\mathfrak{P}$-integral coefficients belonging to Ω. When we replace every coefficient by its residue class (mod $\mathfrak{P}$), we obtain a modular representation $\bar{\mathfrak{Z}}$ whose coefficients belong to the field $\bar{\Omega}$ of residue classes of integers of Ω (mod $\mathfrak{P}$). After replacing $\mathfrak{Z}$ by a similar representation, we may assume that $\bar{\mathfrak{Z}}$ splits into irreducible constituents in $\bar{\Omega}$.

If $\bar{\mathfrak{F}}$ is any irreducible representation of $\mathfrak{G}$ in the algebraically closed extension field of $\bar{\Omega}$, then the traces of all the matrices are sums of modular g^{th} roots of unity. Since $\bar{\Omega}$ has the characteristic p, they are sums of g'^{th} roots of unity, and hence they belong to $\bar{\Omega}$. It follows that $\bar{\mathfrak{F}}$ can be written with coefficients in $\bar{\Omega}$.[12] Any modular representation which is irreducible in $\bar{\Omega}$ is absolutely irreducible. We may set

$$\mathfrak{Z} = \begin{pmatrix} \mathfrak{A}_{11} & \mathfrak{A}_{12} & \cdots \\ \mathfrak{A}_{21} & \mathfrak{A}_{22} & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix}, \tag{1}$$

where $\mathfrak{A}_{\kappa\lambda} \equiv 0 \pmod{\mathfrak{P}}$ for $\kappa < \lambda$, and where the $\mathfrak{A}_{\kappa\kappa} \pmod{\mathfrak{P}}$ are the absolutely irreducible modular constituents of $\mathfrak{Z}$.

2. On the number of p-conjugate characters

The degree of Ω over K is $m = \varphi(p^\alpha) = p^{\alpha-1}(p-1)$; we denote the conjugates of a number ω of Ω by $\omega, \omega', \cdots, \omega^{(m-1)}$, and use the corresponding notation for matrices and representations. Then $\mathfrak{Z}^{(\rho)}$ and $\mathfrak{Z}^{(\sigma)}$ are either non-similar or identical.

Let $\mathfrak{Z}(G)$ be the matrix representing the group element G, and denote by $\beta(G)$ the coefficient in the upper right corner of $\mathfrak{Z}(G)$ and by $\gamma(G)$ the coefficient in the lower left corner of $\mathfrak{Z}(G)$. The fundamental relations of I. Schur[13] for the coefficients of a representation give

$$\sum_{G \text{ in } \mathfrak{G}} \beta(G)^{(\rho)} \gamma(G^{-1})^{(\sigma)} = \eta_{\rho\sigma} g/z, \tag{2}$$

where z is the degree of $\mathfrak{Z}$, and $\eta_{\rho\sigma} = 0$ for $\zeta^{(\rho)} \neq \zeta^{(\sigma)}$, and $\eta_{\rho\sigma} = 1$ for $\zeta^{(\rho)} = \zeta^{(\sigma)}$. Let ω be any integer of Ω, and set

$$\xi_1(G) = \sum_{\rho=0}^{m-1} \omega^{(\rho)} \beta(G)^{(\rho)} = \operatorname{tr}(\omega\beta(G)); \qquad \xi_2(G) = \sum_{\sigma=0}^{m-1} \gamma(G)^{(\sigma)} = \operatorname{tr}(\gamma(G)), \tag{3}$$

where tr $(\cdots)$ denotes the trace of an element of Ω with regard to K. Then (2) yields

$$\sum_G \xi_1(G)\xi_2(G^{-1}) = (g/z) \cdot \sum_{\rho,\sigma} \eta_{\rho\sigma} \omega^{(\rho)}.$$

---

[11] We then have $\mathfrak{p} = \mathfrak{P}^e$, $e = p^{\alpha-1}(p-1)$.

[12] Cf. R. Brauer, Math. Zeitschr. 29, p. 79 (1929), p. 101. The fact used is equivalent to Wedderburn's theorem on finite division rings. The argument shows that $\bar{\mathfrak{F}}$ can be written in the field of the modular g'th roots of unity and, therefore, in the field $\bar{\mathrm{K}}$ of residue classes of integers of K (mod $\mathfrak{p}$).

[13] I. Schur, Sitzungsber. Preuss. Akad. d. Wiss. 1905 p. 406, theorem IV.

If r is the degree of $K(\zeta)$ over K, we have m/r characters $\zeta^{(\rho)}$ which are equal to a fixed $\zeta^{(\mu)}$. Hence

$$\sum_G \xi_1(G)\xi_2(G^{-1}) = \frac{gm}{zr} \operatorname{tr}(\omega). \tag{4}$$

The trace of any $\mathfrak{P}$-integer is divisible by $\mathfrak{p}^{\alpha-1}$.[14] This holds, in particular, for $\xi_1(G)$ and $\xi_2(G)$. For $\omega = 1$, we thus find

$$\frac{gm^2}{rz} \equiv 0 \pmod{\mathfrak{p}^{2\alpha-2}}.$$

Since the left side is rational, $\mathfrak{p}$ may be replaced by p (cf. lemma 1*, (β)). Since $m = \varphi(p^\alpha)$, this gives

THEOREM 1: *If ζ is an irreducible character of degree z, and if r is the number of p-conjugate characters, then $g/(rz)$ is a p-integer for any prime p.*

We next assume that $\mathfrak{Z}$ is reducible (mod $\mathfrak{P}$). Then (1) shows that $\beta(G) \equiv 0$ (mod $\mathfrak{P}$). Set

$$v = p^{\alpha-1}, \qquad \Theta = (1 - \epsilon^v)/(1 - \epsilon) = 1 + \epsilon + \cdots + \epsilon^{v-1}. \tag{5}$$

Since ϵ^v is a primitive p^{th} root of unity, Θ is divisible by $\mathfrak{P}^{v-1}$. If a $\mathfrak{P}$-integer is divisible by $\mathfrak{P}^v$, its trace is divisible by $\mathfrak{p}^\alpha$.[15] For $\omega = \Theta$, we have $\omega\beta(G) \equiv 0$ (mod $\mathfrak{P}^v$) and, therefore, $\xi_1(G) \equiv 0$ (mod $\mathfrak{p}^\alpha$). As before, $\xi_2(G) \equiv 0$ (mod $\mathfrak{p}^{\alpha-1}$). From (5) it follows that $\operatorname{tr}(\Theta) = m$, and (4) now yields

$$gm^2/rz \equiv 0 \pmod{\mathfrak{p}^{2\alpha-1}}.$$

Hence

LEMMA 2: *If ζ in theorem 1 is reducible as a modular character, then the p-integer g/rz is still divisible by p.*

3. Characters with a common modular constituent

We use a similar argument in order to study two irreducible representations $\mathfrak{Z}_1$ and $\mathfrak{Z}_2$ which have a modular constituent in common. We exclude the case when the characters ζ_1 and ζ_2 of $\mathfrak{Z}_1$ and $\mathfrak{Z}_2$ are p-conjugate. The $p^{\alpha\,\text{th}}$ root of unity ϵ may be so chosen that $\Omega = K(\epsilon)$ contains both ζ_1 and ζ_2; let $\mathfrak{H}$ be the Galois group of Ω with regard to K.

Each $\mathfrak{Z}_\mu$ can be written in the form (1), say $\mathfrak{Z}_\mu = (\mathfrak{A}^\mu_{i\lambda})$. For at least one pair of indices i, j, we must have

$$\mathfrak{A}^1_{ii} \equiv \mathfrak{A}^2_{jj} \pmod{\mathfrak{P}}. \tag{6}$$

---

[14] Any $\mathfrak{P}$-integer λ can be written in the form $\lambda = \sum u_\mu \epsilon^\mu$ where $\mu = 0, 1, \cdots, m-1$ and the u_μ are $\mathfrak{p}$-integers of K. We have $\operatorname{tr}(\epsilon^\mu) = m$ if $\mu \equiv 0 \pmod{p^\alpha}$, $\operatorname{tr}(\epsilon^\mu) = -m/(p-1)$, if $\mu \equiv 0 \pmod{p^{\alpha-1}}$, $\mu \not\equiv 0 \pmod{p^\alpha}$, and $\operatorname{tr}(\epsilon^\mu) = 0$ in all other cases. This proves the statement.

[15] It is sufficient to prove this for numbers of the form $(1 - \epsilon^v)\epsilon^\mu$, cf. footnote 14, and here it is evident.

Let $\gamma_1(G)$ be the coefficient in the upper left corner of $\mathfrak{A}_{ii}^1$, and $\gamma_2(G)$ the coefficient in the upper left corner of $\mathfrak{A}_{jj}^2$, and set $\beta(G) = \gamma_1(G) - \gamma_2(G)$ so that

$$\beta(G) = \gamma_1(G) - \gamma_2(G) \equiv 0 \pmod{\mathfrak{P}}. \tag{7}$$

Schur's relations[13] give here

$$\sum_G \gamma_1(G)^{(\rho)}\gamma_1(G^{-1})^{(\sigma)} = \frac{g}{z_1}\eta'_{\rho\sigma}; \qquad \sum_G \gamma_2(G)^{(\rho)}\gamma_2(G^{-1})^{(\sigma)} = \frac{g}{z_2}\eta''_{\rho\sigma};$$
$$\sum_G \gamma_1(G)^{(\rho)}\gamma_2(G^{-1})^{(\sigma)} = 0, \tag{8}$$

where z_μ is the degree of $\mathfrak{Z}_\mu$, and $\eta^\mu_{\rho\sigma} = 0$ or 1 according as $\zeta_\mu^{(\rho)} \neq \zeta_\mu^{(\sigma)}$ or $\zeta_\mu^{(\rho)} = \zeta_\mu^{(\sigma)}$.

The same relations hold, when $\mathfrak{Z}_1 = \mathfrak{Z}_2$, and $\mathfrak{Z}_1$ contains one of its modular constituents more than once. We then have a formula (6) with $i \neq j$, and (7) and (8) are also true.

From (8) it follows that

$$\sum_G \beta(G)^{(\rho)}\beta(G^{-1})^{(\sigma)} = \frac{g}{z_1}\eta'_{\rho\sigma} + \frac{g}{z_2}\eta''_{\rho\sigma}. \tag{9}$$

We set (cf. (3))

$$\xi_1(G) = \operatorname{tr}(\omega\beta(G)), \qquad \xi_2(G) = \operatorname{tr}(\psi\beta(G)), \tag{10}$$

where ω and ψ are two integers of Ω. Then (9) gives

$$\sum_G \xi_1(G)\xi_2(G^{-1}) = \frac{g}{z_1}\sum_{\rho,\sigma}\eta'_{\rho\sigma}\omega^{(\rho)}\psi^{(\sigma)} + \frac{g}{z_2}\sum_{\rho,\sigma}\eta''_{\rho\sigma}\omega^{(\rho)}\psi^{(\sigma)}.$$

When $K(\zeta_\mu)$ corresponds to the subgroup $\mathfrak{L}_\mu$ of the Galois group $\mathfrak{H}$, this can be written in the form

$$\sum_G \xi_1(G)\xi_2(G^{-1}) = u_1\frac{g}{z_1} + u_2\frac{g}{z_2}, \tag{11}$$

$$u_\mu = \sum_{X_1 \equiv X_2 \pmod{\mathfrak{L}_\mu}} \omega^{X_1}\psi^{X_2}, \tag{12}$$

where in (12) the sum is to be extended over all pairs of elements X_1, X_2 of $\mathfrak{H}$, for which $X_1X_2^{-1}$ belongs to $\mathfrak{L}_\mu$.

We first choose $\omega = \Theta$, $\psi = 1$, where Θ is defined in (5). Then $\omega\beta(G) \equiv 0 \pmod{\mathfrak{P}^v}$, and as in §2, the trace $\xi_1(G)$ of $\omega\beta(G)$ is divisible by $\mathfrak{p}^\alpha$. Furthermore, $\xi_2(G)$ is divisible by $\mathfrak{p}^{\alpha-1}$, since it is a trace. If $\alpha = 1$, we may even state $\xi_2(G) \equiv 0 \pmod{\mathfrak{p}}$, because $\beta(G)$ is divisible by $\mathfrak{P}$. Finally, $u_\mu = \operatorname{tr}(\Theta)m/r_\mu = m^2/r_\mu$, since $\mathfrak{L}_\mu$ has the order m/r_μ and $\operatorname{tr}(\Theta) = m$. Then (11) yields

$$gm^2/r_1z_1 + gm^2/r_2z_2 \equiv 0 \pmod{\mathfrak{p}^{2\alpha-1}}.$$

For $\alpha = 1$, this congruence even holds (mod $\mathfrak{p}^2$). If $(r_1, p) = 1$ and $(r_2, p) = 1$, we may choose $\alpha = 1$. As before, $\mathfrak{p}$ can be replaced by p. Since $m = p^{\alpha-1}(p-1)$, we have

LEMMA 3: *Let ζ_1 and ζ_2 be two characters of $\mathfrak{G}$ which are not p-conjugate but which have a modular constituent in common. If ζ_μ has r_μ p-conjugates, and if z_μ is the degree of ζ_μ, then*

$$g/r_1z_1 + g/r_2z_2 \equiv 0 \pmod{p}, \tag{13}$$

$$g/r_1z_1 + g/r_2z_2 \equiv 0 \pmod{p^2}, \quad \textit{if} \quad (r_1, p) = 1 \quad \textit{and} \quad (r_2, p) = 1. \tag{13*}$$

Secondly, we choose ω, ψ such that $\omega \equiv 0, \psi \equiv 0 \pmod{\mathfrak{P}^{v-1}}$. Then $\xi_1(G) \equiv \xi_2(G) \equiv 0 \pmod{p^a}$, and we obtain

LEMMA 4: *Suppose that ζ_1 and ζ_2 satisfy the assumptions of lemma 3. Let ω and ψ be two integers of $K(\epsilon)$ which are divisible by $\mathfrak{P}^{v-1}$, $v = p^{a-1}$. Then*

$$u_1g/z_1 + u_2g/z_2 \equiv 0 \pmod{p^{2a}}, \tag{14}$$

where u_μ is defined by (12).

According to a remark above, the same argument will hold, if $\mathfrak{Z}_1 = \mathfrak{Z}_2$ and $\mathfrak{Z}_1$ contains one of its modular constituents more than once.

LEMMA 5: *If ζ is a character which contains one of its modular constituents more than once, then for odd p,*

$$g/(rz) \equiv 0 \pmod{p^2}, \tag{15}$$

where z is the degree of ζ, and where the number r of p-conjugates of ζ is assumed to be prime to p.

4. On representations for which $rz \equiv 0 \pmod{p^a}$

We now prove

THEOREM 2: *Let ζ be an irreducible character of degree z which has r distinct p-conjugates. Then rz can be divisible by p^a only if z is divisible by p^a, i.e. when ζ is of the highest kind.*

PROOF: Assume that $rz \equiv 0 \pmod{p^a}$, $z \not\equiv 0 \pmod{p^a}$. Since the corresponding representation $\mathfrak{Z}$ is not of the highest kind, the block B of $\zeta = \zeta_1$ cannot consist of the family of ζ_1 only.[16] Hence we may find a character ζ_2 in B which is not p-conjugate to ζ_1, but has a modular constituent in common with ζ_1. Then (13) shows that also $r_2z_2 \equiv 0 \pmod{p^a}$. On the other hand, $z_2 \not\equiv 0 \pmod{p^a}$, because B is not of the highest kind. Consequently, ζ_2 satisfies the same assumptions as ζ_1. On considering chains of characters of B such that any two consecutive terms have a modular character in common, we conclude that every character ζ_ν of B satisfies the conditions $r_\nu z_\nu \equiv 0 \pmod{p^a}$, $z_\nu \not\equiv 0 \pmod{p^a}$, where z_ν again is the degree of ζ_ν and r_ν the number of p-conjugates.

It now follows from lemma 2 that all the ζ_ν of B are modular-irreducible, and this implies that they all are equal when considered as modular characters. To the block $B = B_\lambda$ there corresponds a part D_λ[17] of the matrix D of the decompo-

[16] R. Brauer, *On the connection between the ordinary and the modular characters of groups of finite order*, Ann. of Math., **42**, pp. 926–935, 1941, theorem 8.

[17] Cf. BN, §9, (28).

sition numbers of $\mathfrak{G}$. Our results show that the matrix D_λ consists of one column only, and all the coefficients are 1. All the ζ_ν have the same degree z, and if $z = p^\rho z'$ with $(z', p) = 1$, then all the r_ν are divisible by $p^{a-\rho}$. The number of rows in D_λ is equal to the sum of all r_ν belonging to the different families of B, and hence this number is larger than $p^{a-\rho}$.

There corresponds to B a part $C_\lambda = D_\lambda' D_\lambda$ of the matrix C of the Cartan invariants.[17] This C_λ is of degree 1; its only coefficient is equal to the number of rows of D_λ, and therefore is larger than $p^{a-\rho}$. On the other hand, the block B_λ is of type ρ, and then C_λ has $p^{a-\rho}$ as an elementary divisor,[18] which means that $C_\lambda = (p^{a-\rho})$. This gives a contradiction, and hence the theorem is proved.

5. Evaluation of the numbers u_μ in lemma 4

Assume that $\alpha \geqq 2$. Let $\mathfrak{o}$ be a system of $p^{\alpha-1} = v$ rational integers ρ_1, $\rho_2, \cdots$ which form a complete residue system (mod $p^{\alpha-1}$), $\alpha \geqq 2$, and let $\mathfrak{d}$: $\sigma_1, \sigma_2, \cdots$ be a second system with the same properties. We set

$$\omega = \Sigma \epsilon^\rho, \qquad \psi = \Sigma \epsilon^{-\sigma},$$

where ρ ranges over the values of $\mathfrak{o}$, and σ over those of $\mathfrak{d}$. If $\rho \equiv \tau \pmod{v}$, then $\epsilon^\rho \equiv \epsilon^\tau \pmod{\mathfrak{P}^v}$ when again $v = p^{\alpha-1}$. It follows that ω and ψ both are congruent to the number Θ in (5) and hence both are divisible by $\mathfrak{P}^{v-1}$ as assumed in lemma 4. Now (12) becomes

$$u_\mu = \sum_{X_1 \equiv X_2 \pmod{\mathfrak{L}_\mu}} \sum_\rho \sum_\sigma \epsilon^{\rho X_1 - \sigma X_2} \qquad (X_1, X_2 \text{ in } \mathfrak{H}).^{19} \tag{16}$$

If ϵ^λ appears, all its conjugates appear with the same multiplicity. Consequently, if A_μ' terms of (16) are equal to 1, and B_μ' terms are primitive p^{th} roots of unity, we have $u_\mu = A_\mu' - B_\mu'/(p-1)$. The term on the right-hand side of (16) is 1, if $\rho X_1 \equiv \sigma X_2 \pmod{p^\alpha}$, i.e. if $\rho X_1 X_2^{-1} \equiv \sigma \pmod{p^\alpha}$. Similarly, the term is a primitive p^{th} root of unity, if this congruence holds (mod $p^{\alpha-1}$) but not (mod p^α). Now $X_1 X_2^{-1}$ is an element of $\mathfrak{L}_\mu$. Let A_μ denote the number of pairs (ρ, L), ρ in $\mathfrak{o}$, L in $\mathfrak{L}_\mu$, for which ρL is congruent to one of the numbers σ (mod p^α), and B_μ the number of pairs, for which a congruence $\rho L \equiv \sigma$ (σ in $\mathfrak{d}$) holds (mod $p^{\alpha-1}$) but not (mod p^α). Then A_μ' and B_μ' are obtained from A_μ and B_μ by multiplying the latter by the order $\varphi(p^\alpha) = v(p-1)$ of $\mathfrak{H}$, and hence

$$u_\mu = v(p-1)A_\mu - vB_\mu . \tag{17}$$

---

[18] In BN, §17, it was proved that at least one elementary divisor corresponding to a block of type α is divisible by $p^{a-\alpha}$. Actually, exactly one elementary divisor is equal to $p^{a-\alpha}$ while the other elementary divisors are smaller than $p^{a-\alpha}$. This can be proved by a generalization of the method of BN, §21; cf. below section 8 of this paper where the same method is used (for $p = 2$).

[19] We may understand ρX_1 and σX_2 as rational integers (mod p^α).

Since every ρL is congruent to one of the σ (mod $p^{\alpha-1}$), and since $\mathfrak{L}_\mu$ is of the order $l_\mu = v(p-1)/r_\mu$, we have $A_\mu + B_\mu = v^2(p-1)/r_\mu$. Then (17) becomes $u_\mu = vpA_\mu - v^3(p-1)/r_\mu$ and (14) takes the form

$$\frac{p^\alpha g}{z_1} A_1 + \frac{p^\alpha g}{z_2} A_2 - \left(\frac{p^{3\alpha-3}(p-1)g}{r_1 z_1} + \frac{p^{3\alpha-3}(p-1)g}{r_2 z_2} \right) \equiv 0 \pmod{p^{2\alpha}}.$$

According to (13), the last two terms are divisible by $p^{3\alpha-2}$, and since $3\alpha - 2 \geqq 2\alpha$, can be neglected. We then have

$$gA_1/z_1 + gA_2/z_2 \equiv 0 \pmod{p^\alpha}. \tag{18}$$

If p is odd, we shall need the value of A_μ, only if r_μ is prime to p. The field $K(\zeta_\mu)$ then is contained in the field $K(\epsilon_p)$ where ϵ_p is a primitive p^{th} root of unity. If $\rho \not\equiv 0 \pmod{v}$, ($\rho$ in $\mathfrak{o}$), then for every L in $\mathfrak{L}_\mu$ the number $\epsilon^{\rho L}$ is conjugate to ϵ^ρ with regard to $K(\zeta_\mu)$. If ϵ^τ appears in the form $\epsilon^{\rho L}$, so do

$$\epsilon^\tau, \epsilon^\tau \epsilon_p, \cdots, \epsilon^\tau \epsilon_p^{p-1}$$

and each of them appears the same number of times. Exactly one of these quantities is of the form ϵ^σ, (σ in $\mathfrak{d}$). Hence, for a fixed ρ, one in every p of the elements L of $\mathfrak{L}_\mu$ satisfies the condition that ρL belongs to $\mathfrak{d}$. For $\rho \equiv 0 \pmod{v}$, the number of elements L of $\mathfrak{L}_\mu$, for which $\epsilon^{\rho L}$ has a fixed value, is divisible by $p^{\alpha-1}$. Hence, in the case $(r_\mu, p) = 1$, we have

$$A_\mu \equiv l_\mu(v-1)/p \equiv v(p-1)(v-1)/r_\mu p \equiv p^{\alpha-2}/r_\mu \pmod{p^{\alpha-1}}. \tag{19}$$

The congruences (18) and (19) enable us to prove the following lemma:

LEMMA 6: *Let p be odd. If ζ_1 and ζ_2 satisfy the assumptions of lemma 3, and if $z_1 \equiv 0 \pmod{p^{\alpha-1}}$, then $r_2 \not\equiv 0 \pmod{p}$.*

PROOF: From theorem 2, it follows that $r_1 \not\equiv 0 \pmod{p}$. Assume that $r_2 \equiv 0 \pmod{p}$. If we choose for α the smallest value for which ζ_1 and ζ_2 belong to the field $K(\epsilon)$ obtained from K by adjoining a p^{ath} root of unity ϵ, then r_2 must be divisible by $p^{\alpha-1}$. Theorem 2 shows that we have $z_2 \not\equiv 0 \pmod{p^{\alpha-\alpha+1}}$ and, therefore, the second term in (18) is divisible by p^α. Since $(r_1, p) = 1$, we may use (19) for $\mu = 1$ and find that the first term in (18) is divisible by $p^{\alpha-1}$ only, which gives a contradiction. Hence $r_2 \not\equiv 0 \pmod{p}$ as was stated.

For $p = 2$, we choose $\alpha \geqq 3$. Here, r_μ is a power of 2. If $r_\mu = 2^j$, $1 \leqq j \leqq \alpha - 2$, the field $K(\epsilon)$ has exactly three subfields of degree 2^j over K. One of them, say $\Gamma_j^{(0)}$, is obtained by adjoining a $2^{j+1\text{th}}$ root of unity to K. We denote the other two by $\Gamma_j^{(1)}$ and $\Gamma_j^{(2)}$, and set $K(\epsilon) = \Gamma_{\alpha-1}^{(0)}$, $K = \Gamma_0^{(1)}$. Then, by an argument similar to the one used for odd p, we may prove

$$A_\mu \equiv 0 \pmod{2^{\alpha-1-j}}, \quad \text{if} \quad K(\zeta_\mu) = \Gamma_j^{(0)}. \tag{20}$$

$$A_\mu \equiv 2^{\alpha-2-j} \pmod{2^{\alpha-1-j}}, \quad \text{if} \quad K(\zeta_\mu) = \Gamma_j^{(1)} \quad \text{or} \quad \Gamma_j^{(2)}. \tag{21}$$

LEMMA 7: *Let $p = 2$, and suppose that ζ_1 and ζ_2 satisfy the assumptions of lemma 3. If $r_1 z_1 \equiv 0 \pmod{2^{\alpha-1}}$ and if $K(\zeta_1)$ is not of the form $K(\tilde{\epsilon})$ where $\tilde{\epsilon}$ is a primitive*

sition numbers of $\mathfrak{G}$. Our results show that the matrix D_λ consists of one column only, and all the coefficients are 1. All the ζ_ν have the same degree z, and if $z = p^\rho z'$ with $(z', p) = 1$, then all the r_ν are divisible by $p^{a-\rho}$. The number of rows in D_λ is equal to the sum of all r_ν belonging to the different families of B, and hence this number is larger than $p^{a-\rho}$.

There corresponds to B a part $C_\lambda = D_\lambda' D_\lambda$ of the matrix C of the Cartan invariants.[17] This C_λ is of degree 1; its only coefficient is equal to the number of rows of D_λ, and therefore is larger than $p^{a-\rho}$. On the other hand, the block B_λ is of type ρ, and then C_λ has $p^{a-\rho}$ as an elementary divisor,[18] which means that $C_\lambda = (p^{a-\rho})$. This gives a contradiction, and hence the theorem is proved.

5. Evaluation of the numbers u_μ in lemma 4

Assume that $\alpha \geqq 2$. Let $\mathfrak{o}$ be a system of $p^{\alpha-1} = v$ rational integers ρ_1, $\rho_2, \cdots$ which form a complete residue system (mod $p^{\alpha-1}$), $\alpha \geqq 2$, and let $\mathfrak{d}$: $\sigma_1, \sigma_2, \cdots$ be a second system with the same properties. We set

$$\omega = \Sigma \epsilon^\rho, \qquad \psi = \Sigma \epsilon^{-\sigma},$$

where ρ ranges over the values of $\mathfrak{o}$, and σ over those of $\mathfrak{d}$. If $\rho \equiv \tau \pmod{v}$, then $\epsilon^\rho \equiv \epsilon^\tau \pmod{\mathfrak{P}^v}$ when again $v = p^{\alpha-1}$. It follows that ω and ψ both are congruent to the number Θ in (5) and hence both are divisible by $\mathfrak{P}^{v-1}$ as assumed in lemma 4. Now (12) becomes

$$u_\mu = \sum_{X_1 \equiv X_2 \pmod{\mathfrak{L}_\mu}} \sum_\rho \sum_\sigma \epsilon^{\rho X_1 - \sigma X_2} \qquad (X_1, X_2 \text{ in } \mathfrak{H}).^{19} \tag{16}$$

If ϵ^λ appears, all its conjugates appear with the same multiplicity. Consequently, if A_μ' terms of (16) are equal to 1, and B_μ' terms are primitive p^{th} roots of unity, we have $u_\mu = A_\mu' - B_\mu'/(p-1)$. The term on the right-hand side of (16) is 1, if $\rho X_1 \equiv \sigma X_2 \pmod{p^\alpha}$, i.e. if $\rho X_1 X_2^{-1} \equiv \sigma \pmod{p^\alpha}$. Similarly, the term is a primitive p^{th} root of unity, if this congruence holds (mod $p^{\alpha-1}$) but not (mod p^α). Now $X_1 X_2^{-1}$ is an element of $\mathfrak{L}_\mu$. Let A_μ denote the number of pairs (ρ, L), ρ in $\mathfrak{o}$, L in $\mathfrak{L}_\mu$, for which ρL is congruent to one of the numbers $\sigma \pmod{p^\alpha}$, and B_μ the number of pairs, for which a congruence $\rho L \equiv \sigma$ (σ in $\mathfrak{d}$) holds (mod $p^{\alpha-1}$) but not (mod p^α). Then A_μ' and B_μ' are obtained from A_μ and B_μ by multiplying the latter by the order $\varphi(p^\alpha) = v(p-1)$ of $\mathfrak{H}$, and hence

$$u_\mu = v(p-1)A_\mu - vB_\mu . \tag{17}$$

---

[18] In BN, §17, it was proved that at least one elementary divisor corresponding to a block of type α is divisible by $p^{a-\alpha}$. Actually, exactly one elementary divisor is equal to $p^{a-\alpha}$ while the other elementary divisors are smaller than $p^{a-\alpha}$. This can be proved by a generalization of the method of BN, §21; cf. below section 8 of this paper where the same method is used (for $p = 2$).

[19] We may understand ρX_1 and σX_2 as rational integers (mod p^α).

Since every ρL is congruent to one of the $\sigma \pmod{p^{\alpha-1}}$, and since $\mathfrak{L}_\mu$ is of the order $l_\mu = v(p-1)/r_\mu$, we have $A_\mu + B_\mu = v^2(p-1)/r_\mu$. Then (17) becomes $u_\mu = vpA_\mu - v^3(p-1)/r_\mu$ and (14) takes the form

$$\frac{p^\alpha g}{z_1} A_1 + \frac{p^\alpha g}{z_2} A_2 - \left(\frac{p^{3\alpha-3}(p-1)g}{r_1 z_1} + \frac{p^{3\alpha-3}(p-1)g}{r_2 z_2} \right) \equiv 0 \pmod{p^{2\alpha}}.$$

According to (13), the last two terms are divisible by $p^{3\alpha-2}$, and since $3\alpha - 2 \geqq 2\alpha$, can be neglected. We then have

$$gA_1/z_1 + gA_2/z_2 \equiv 0 \pmod{p^\alpha}. \tag{18}$$

If p is odd, we shall need the value of A_μ, only if r_μ is prime to p. The field $K(\zeta_\mu)$ then is contained in the field $K(\epsilon_p)$ where ϵ_p is a primitive p^{th} root of unity. If $\rho \not\equiv 0 \pmod{v}$, ($\rho$ in $\mathfrak{o}$), then for every L in $\mathfrak{L}_\mu$ the number $\epsilon^{\rho L}$ is conjugate to ϵ^ρ with regard to $K(\zeta_\mu)$. If ϵ^τ appears in the form $\epsilon^{\rho L}$, so do

$$\epsilon^\tau, \ \epsilon^\tau \epsilon_p, \ \cdots, \ \epsilon^\tau \epsilon_p^{p-1}$$

and each of them appears the same number of times. Exactly one of these quantities is of the form ϵ^σ, (σ in $\mathfrak{d}$). Hence, for a fixed ρ, one in every p of the elements L of $\mathfrak{L}_\mu$ satisfies the condition that ρL belongs to $\mathfrak{d}$. For $\rho \equiv 0 \pmod{v}$, the number of elements L of $\mathfrak{L}_\mu$, for which $\epsilon^{\rho L}$ has a fixed value, is divisible by $p^{\alpha-1}$. Hence, in the case $(r_\mu, p) = 1$, we have

$$A_\mu \equiv l_\mu(v-1)/p \equiv v(p-1)(v-1)/r_\mu p \equiv p^{\alpha-2}/r_\mu \pmod{p^{\alpha-1}}. \tag{19}$$

The congruences (18) and (19) enable us to prove the following lemma:

LEMMA 6: *Let p be odd. If ζ_1 and ζ_2 satisfy the assumptions of lemma 3, and if $z_1 \equiv 0 \pmod{p^{\alpha-1}}$, then $r_2 \not\equiv 0 \pmod{p}$.*

PROOF: From theorem 2, it follows that $r_1 \not\equiv 0 \pmod{p}$. Assume that $r_2 \equiv 0 \pmod{p}$. If we choose for α the smallest value for which ζ_1 and ζ_2 belong to the field $K(\epsilon)$ obtained from K by adjoining a $p^{\alpha\text{th}}$ root of unity ϵ, then r_2 must be divisible by $p^{\alpha-1}$. Theorem 2 shows that we have $z_2 \not\equiv 0 \pmod{p^{\alpha-\alpha+1}}$ and, therefore, the second term in (18) is divisible by p^α. Since $(r_1, p) = 1$, we may use (19) for $\mu = 1$ and find that the first term in (18) is divisible by $p^{\alpha-1}$ only, which gives a contradiction. Hence $r_2 \not\equiv 0 \pmod{p}$ as was stated.

For $p = 2$, we choose $\alpha \geqq 3$. Here, r_μ is a power of 2. If $r_\mu = 2^j$, $1 \leqq j \leqq \alpha - 2$, the field $K(\epsilon)$ has exactly three subfields of degree 2^j over K. One of them, say $\Gamma_j^{(0)}$, is obtained by adjoining a $2^{j+1\text{th}}$ root of unity to K. We denote the other two by $\Gamma_j^{(1)}$ and $\Gamma_j^{(2)}$, and set $K(\epsilon) = \Gamma_{\alpha-1}^{(0)}$, $K = \Gamma_0^{(1)}$. Then, by an argument similar to the one used for odd p, we may prove

$$A_\mu \equiv 0 \pmod{2^{\alpha-1-j}}, \quad \text{if} \quad K(\zeta_\mu) = \Gamma_j^{(0)}. \tag{20}$$

$$A_\mu \equiv 2^{\alpha-2-j} \pmod{2^{\alpha-1-j}}, \quad \text{if} \quad K(\zeta_\mu) = \Gamma_j^{(1)} \quad \text{or} \quad \Gamma_j^{(2)}. \tag{21}$$

LEMMA 7: *Let $p = 2$, and suppose that ζ_1 and ζ_2 satisfy the assumptions of lemma 3. If $r_1 z_1 \equiv 0 \pmod{2^{\alpha-1}}$ and if $K(\zeta_1)$ is not of the form $K(\bar{\epsilon})$ where $\bar{\epsilon}$ is a primitive*

$2^{j\text{th}}$ *root of unity with* $j \geqq 2$, *then* $r_2 z_2 \equiv 0 \pmod{2^{a-1}}$, *and* $K(\zeta_2)$ *is not of the form* $K(\epsilon)$.

PROOF: From the assumptions, theorem 2, and (21), it follows that $A_1 g/z_1$ is divisible by 2^{a-1}, but not by 2^a. According to (18), $A_2 g/z_2$ then is also divisible by 2^{a-1} but not by 2^a. If we set $r_2 = 2^j$ and use (20), we see that $K(\zeta_2)$ cannot be one of the fields $\Gamma_j^{(0)}$, $(j = 1, 2, \cdots)$. If $K(\zeta_2) = \Gamma_j^{(1)}$ and $K(\zeta_2) = \Gamma_j^{(2)}$ we use (21) and conclude that z_2 is divisible by 2^{a-1-j}, which implies $r_2 z_2 \equiv 0 \pmod{2^{a-1}}$.

6. Representations of a degree which is divisible by p^{a-1}

Let ζ be an irreducible character of degree $z \equiv 0 \pmod{p^{a-1}}$. If z is divisible by p^a, then ζ is of the highest kind,[20] and we will exclude this case in the following. We shall first assume that p is odd.[21] The block B to which $\zeta = \zeta_1$ belongs is not of the highest kind. Then B does not consist of the family of ζ only,[22] and we may therefore find a character ζ_2 which is not p-conjugate to $\zeta = \zeta_1$ but has a modular constituent in common with ζ_1 (cf. the analogous argument in §4). If z_μ is the degree of ζ_μ, and r_μ the number of p-conjugate characters, then theorem 2 and lemma 6 show that $r_1 \not\equiv 0$, $r_2 \not\equiv 0 \pmod{p}$. In (13*), lemma 3, the first term on the left side is divisible by p but not by p^2. The same must hold for the second term, and hence we must have $z_2 \equiv 0 \pmod{p^{a-1}}$, so that ζ_2 satisfies the same assumption as ζ. Continuing in this manner we see that the degree of every character of B is divisible by p^{a-1}. This gives

THEOREM 3: *If the degree of one character of a block* B *is divisible by* p^{a-1}, *the same is true for all characters of* B.

In the notation of BN, the block B is of the type $a - 1$. If ζ_1 and ζ_2 are again two characters of such a block B which are not p-conjugate but have a modular constituent in common, then on multiplying (13*) by $r_1 r_2 z_1 z_2/g \equiv 0 \pmod{p^{a-2}}$ the congruence $r_1 z_1 + r_2 z_2 \equiv 0 \pmod{p^a}$ is obtained. Now any two characters ζ_κ, ζ_μ can be joined by a chain $\zeta_\kappa, \zeta_\rho, \cdots, \zeta_\sigma, \zeta_\mu$ of characters, such that any two consecutive terms of the chain have a modular character in common without being p-conjugate. It follows that for any ζ_μ of B, we have $r_\mu z_\mu \equiv \pm r_1 z_1 \pmod{p^a}$. Since $r_1 z_1 \not\equiv 0 \pmod{p^a}$ and p is odd, for each μ only one of the signs can be used. Hence

THEOREM 4: *The characters of a block* B *of type* $a - 1$ *can be distributed in two subsets* B′ *and* B″ *such that every character belongs to exactly one of these subsets. If* z_μ *is the degree of a character of* B, *and* r_μ *the number of* p*-conjugates, we may find a rational integer* N *such that* $r_\mu z_\mu \equiv N \pmod{p^a}$ *for all* ζ_μ *of* B′ *and* $r_\mu z_\mu \equiv -N \pmod{p^a}$ *for all* ζ_μ *in* B″. *If* ζ_μ *and* ζ_ν *are not* p*-conjugate and belong both to the same subset* $B^{(\tau)}$, *then they have no modular constituent in common.*

Finally, lemma 5 can be applied and yields

---

[20] Cf. BN, Part II.

[21] The case $p = 2$ will be treated in §8.

[22] Cf. footnote 16.

THEOREM 5: *If ζ belongs to a block* B *of type* $a - 1$, *then it contains each of its modular constituents only with the multiplicity* 1.

7. The matrix D_λ corresponding to a block B of type $a - 1$

NOTATION: $B = B_\lambda$ is a block of type $a - 1$ consisting of the ordinary characters $\zeta_1, \zeta_2, \cdots, \zeta_x$ and the modular characters $\varphi_1, \varphi_2, \cdots, \varphi_y$. The degree of ζ_μ is z_μ, the number of p-conjugates r_μ. There are w families of p-conjugate characters ζ_μ in B, and the ζ_μ are arranged so that $\zeta_1, \zeta_2, \cdots, \zeta_w$ lie in different families, while the characters of the subset B'_λ (theorem 4) come before the characters of the other subset B''_λ. For p-regular elements G of $\mathfrak{G}$, we have

$$\zeta_\mu(G) = \sum d_{\mu\nu}\varphi_\nu(G), \qquad (\mu = 1, 2, \cdots, x). \tag{22}$$

We set $D_\lambda = (d_{\mu\nu})$, $(\mu = 1, 2, \cdots, x;\ \nu = 1, 2, \cdots, y)$ and denote the matrix occupying the first w rows of D_λ by $\tilde{D}_\lambda$, i.e. $\tilde{D}_\lambda = (d_{\mu\nu})$, $(\mu = 1, 2, \cdots, w;\ \nu = 1, 2, \cdots, y)$. The matrix D_λ has the same rows as $\tilde{D}_\lambda$, the μ^{th} row of $\tilde{D}_\lambda$ appearing r_μ times in D_λ, $(\mu = 1, 2, \cdots, w)$.

The theorems 4 and 5 give at once

THEOREM 6: *Every column of $\tilde{D}_\lambda$ contains exactly two coefficients* 1, *one in a row corresponding to a character ζ_λ of B'_λ, and the other in a row corresponding to a character ζ_μ of B''_λ. All the other coefficients in the column are zero.*

We also can prove easily

THEOREM 7: *Let $\xi = (\xi_1, \xi_2, \cdots, \xi_w)$ be a row with w elements such that $\xi\tilde{D}_\lambda = 0$. Then ξ is a scalar multiple of the row $(\delta_1, \delta_2, \cdots, \delta_w)$, where $\delta_\mu = 1$ if ζ_μ belongs to B'_λ, and $\delta_\mu = -1$, if ζ_μ belongs to B''_λ.*

PROOF: Two characters $\zeta_1, \zeta_\mu (\mu = 1, 2, \cdots, w)$ of B can be joined by a chain of characters ζ_ρ, $(\rho = 1, 2, \cdots, w)$, such that two consecutive characters ζ_i, ζ_j of the chain have a modular character in common. Then ζ_i, ζ_j belong to different subsets B'_λ, B''_λ (theorem 4). There must be a column of $\tilde{D}_\lambda$ which contains the coefficients 1 in the i^{th} and the j^{th} row. Then $\xi\tilde{D}_\lambda = 0$ implies $\xi_i = -\xi_j$, and we obtain successively $\xi_\mu = \delta_\mu\xi_1$, $(\xi_1, \cdots, \xi_w) = \delta_1\xi_1(\delta_1, \cdots, \delta_w)$ which proves the theorem.

The rank of $\tilde{D}_\lambda$ is y.[23] Hence we have

COROLLARY 1: *The number w of families in* B *and the number y of modular characters of* B *are connected by the formula*

$$w = y + 1. \tag{23}$$

We further state

COROLLARY 2: *If $\tilde{\zeta}_\mu$ is the sum of the r_μ characters which are p-conjugate to $\zeta_\mu (\mu = 1, 2, \cdots, w)$, then $\delta_1\tilde{\zeta}_1(G) = \delta_\mu\tilde{\zeta}_\mu(G)$ for any p-singular element of G of $\mathfrak{G}$.*

Indeed, the numbers $\tilde{\zeta}_1(G), \tilde{\zeta}_2(G), \cdots, \tilde{\zeta}_w(G)$ form a solution of $\xi\tilde{D}_\lambda = 0$, (cf. equation (11) of the paper mentioned in footnote 16). Similarly, equation (6) of the same paper gives

---

[23] Cf. BN, §9.

COROLLARY 3: *Let $d^i_{\mu\nu}$ be the higher decomposition numbers, $i > 0$, and set $\tilde{d}^i_{\mu\nu} = \sum d^i_{\rho\nu}$ where the sum extends over all the r_μ values ρ for which ζ_ρ is p-conjugate to ζ_μ. Then $\delta_1 \tilde{d}^i_{1\nu} = \delta_\mu \tilde{d}^i_{\mu\nu}$.*

For p-regular elements G, the formula (22), together with $(\delta_1, \cdots, \delta_w)\tilde{D}_\lambda = 0$ gives

COROLLARY 4: *For p-regular elements G of $\mathfrak{G}$, we have*

$$\sum_{\mu=1}^{w} \delta_\mu \zeta_\mu(G) = 0. \tag{24}$$

In particular, for $G = 1$, this gives

$$\sum_{\mu=1}^{w} \delta_\mu z_\mu = 0. \tag{25}$$

The matrix $\tilde{D}_\lambda$ has w minors of degree $y = w - 1$, and these, if properly arranged and taken with suitable signs, form a solution of $\xi\tilde{D}_\lambda = 0$. Let Δ be a fixed minor. Then, according to theorem 7, every minor has the value $\pm\Delta$. The minor obtained by removing the μ^{th} column of $\tilde{D}_\lambda$ will appear $r_1 r_2 \cdots r_w / r_\mu$ times as a minor of D_λ. We now form the determinant of $C_\lambda = D'_\lambda D_\lambda$. On the one hand, this determinant has the value p.[24] On the other hand, its value is the sum of the squares of all the minors of D_λ. This gives

$$\sum_{\mu=1}^{w} \Delta^2 (r_1 r_2 \cdots r_w / r_\mu) = p.$$

Consequently, we must have $\Delta = \pm 1$, and we find

COROLLARY 5: *The numbers r_μ satisfy the equation*

$$r_1 r_2 \cdots r_w \left(\frac{1}{r_1} + \frac{1}{r_2} + \cdots + \frac{1}{r_w} \right) = p. \tag{26}$$

The numbers r_μ are divisors of $p - 1$; (26) shows that any two of them are relatively prime.

We associate w distinct points $P_1, P_2, \cdots, P_w$ with the characters $\zeta_1, \zeta_2, \cdots, \zeta_w$; we join P_i and P_j when ζ_i and ζ_j have a modular character in common. The linear graph thus obtained characterizes the matrix $\tilde{D}_\lambda$ completely (apart from the arrangements of the columns). We prove

COROLLARY 6: *The linear graph associated with a block* B *of type $a - 1$ is a tree.*

PROOF: This follows from the facts that the graph is connected and has one more vertex than it has edges.

Each point P_μ which lies only on one of the edges corresponds to a row of $\tilde{D}_\lambda$ which contains only one coefficient 1, and this means that ζ_μ is a modular-irreducible character.

---

[24] The determinant of C is a power of p, cf. R. Brauer, *On the Cartan invariants of groups of finite order*, Ann. of Math. **42**, p. 53, 1941. It follows from the result of footnote 18, that the exponent must be 1.

COROLLARY 7: *Every block* B *of type* $a - 1$ *contains at least two characters* ζ_μ *which are not* p*-conjugate and which are modular-irreducible characters.*

8. The case $p = 2$

For $p = 2$, several of the preceding proofs do not hold. We shall now treat this case, using a different kind of argument.

Let ζ be an irreducible character of degree z where $z \equiv 0 \pmod{2^{a-1}}$ but $z \not\equiv 0 \pmod{2^a}$; here a is again the highest exponent for which $2^a = p^a$ divides g. From theorem 2, it follows that ζ is p-conjugate only to itself, i.e. $r = 1$, and ζ belongs to K. Since K is not of the form $K(\tilde{\epsilon})$ where $\tilde{\epsilon}$ is a primitive $2^{j\text{th}}$ root of unity with $j \geqq 2$, we may apply lemma 7, §5. It follows that if the character ζ_μ has a modular constituent in common with ζ_1 then $r_\mu z_\mu \equiv 0 \pmod{2^{a-1}}$, (where z_μ denotes again the degree of ζ_μ, and r_μ the number of 2-conjugates). Further, $K(\zeta_\mu)$ also is not of the form $K(\tilde{\epsilon})$. Therefore, lemma 7 again can be applied to ζ_μ and any character which has a modular constituent in common with ζ_μ without being p-conjugate to ζ_μ. We finally see that the congruence $r_\mu z_\mu \equiv 0 \pmod{2^{a-1}}$ is true for any character ζ_μ of the block B to which ζ belongs.

Let $\zeta_1, \zeta_2, \cdots, \zeta_x$ be the ordinary characters of B, and $\varphi_1, \varphi_2, \cdots, \varphi_y$ the modular characters of B, and set

$$z_\mu = 2^{r_\mu} z'_\mu, \qquad (z'_\mu, 2) = 1. \tag{27}$$

It now follows easily that

$$r_\mu = 2^{a-1-r_\mu}. \tag{28}$$

Suppose that $r_1 = m$ is the smallest of the numbers $r_1, r_2, \cdots, r_x$.

We consider the sums

$$S_{\mu\nu} = \sum_{\kappa=1}^{k_0} g_\kappa \zeta_\mu(G_\kappa)\zeta_\nu(G_\kappa^{-1}), \tag{29}$$

where G_κ ranges over a system of representatives for the 2-regular classes of $\mathfrak{G}$, and where g_κ denotes the number of elements in the class of G_κ. The value of $S_{\mu\nu}$ is a rational integer. Since $g_\kappa \zeta_\mu(G_\kappa)/z_\mu = \omega_{\mu\kappa}$ is an algebraic integer, it follows easily that $S_{\mu\nu}$ is divisible by z_μ and hence by 2^m. On the other hand, we have[25] $\omega_{\mu\kappa} \equiv \omega_{1\kappa}$ modulo a prime ideal divisor of 2. Hence

$$S_{\mu\nu} = z_\mu \sum_\kappa \omega_{\mu\kappa}\zeta_\nu(G_\kappa^{-1}) \equiv z_\mu \sum_\kappa \omega_{1\kappa}\zeta_\nu(G_\kappa^{-1}) = \frac{z_\mu}{z_1} S_{1\nu} \pmod{2^{r_\mu+1}}.$$

On applying the same argument to $S_{1\nu} = S_{\nu 1}$, we arrive at

$$S_{\mu\nu} \equiv 0 \pmod{2^m}, \tag{30}$$

$$S_{\mu\nu} = S_{\nu\mu} \equiv \frac{z_\mu z_\nu}{z_1^2} S_{11} \pmod{2^{r_\mu+1}}. \tag{31}$$

[25] BN, §9.

If the block B is the λ^{th} of the blocks $B_1, B_2, \cdots$ of $\mathfrak{G}$, and if $D_\lambda = (d_{\mu\nu})$, $C_\lambda = (c_{\mu\nu})$ are the parts[26] of the matrices D and C corresponding to the block B_λ, then (22) gives for the matrix $(S_{\mu\nu})_{\mu,\nu}$

$$(S_{\mu\nu})_{\mu,\nu} = \left(\sum_{\kappa} \sum_{\rho,\sigma=1}^{y} g_\kappa d_{\mu\rho}\varphi_\rho(G_\kappa)\, d_{\nu\sigma}\varphi_\sigma(G_\kappa^{-1})\right)_{\mu,\nu} = D_\lambda(gC_\lambda^{-1})D_\lambda'.^{27}$$

The relation (31) can be written in the form

$$(S_{\mu\nu})_{\mu,\nu} = D_\lambda(gC_\lambda^{-1})D_\lambda' = S_{11}\left(\frac{z_\mu z_\nu}{z_1^2}\right)_{\mu,\nu} + H, \tag{32}$$

where H is a matrix in which every coefficient in both the μ^{th} row and μ^{th} column is divisible by $2^{r_\mu+1}$. The first matrix on the right side in (32) has the rank 1; every coefficient in both the μ^{th} row and μ^{th} column is divisible by 2^{r_μ}.

We now proceed to discuss the elementary divisors of $(S_{\mu\nu})$, choosing the ring of all 2-integers as the underlying domain. If C_λ has the elementary divisors $e_1, e_2, \cdots, e_y$ with $e_\nu = 2^{\epsilon_\nu}$, then

$$|C_\lambda| = e_1 e_2 \cdots e_y = 2^{\epsilon_1+\epsilon_2+\cdots+\epsilon_y}.^{28}$$

The elementary divisors of gC_λ^{-1} are $2^{a-\epsilon_y}, \cdots, 2^{a-\epsilon_2}, 2^{a-\epsilon_1}$, and, since the columns of D_λ are linearly independent (mod 2),[29] the elementary divisors of $D_\lambda(gC^{-1})D_\lambda'$ are given by

$$2^{a-\epsilon_y}, 2^{a-\epsilon_{y-1}}, \cdots, 2^{a-\epsilon_1}, \quad 0, \cdots, 0. \tag{33}$$

On the other hand, the right side of (32) can also be used for a discussion of these elementary divisors. If M is a minor of degree j of $(S_{\mu\nu})$ involving the rows $\mu_1, \mu_2, \cdots, \mu_j$, then a simple computation shows that M is divisible by 2 to the exponent $\sum \tau_\mu + j - 1$, $(\mu = \mu_1, \mu_2, \cdots, \mu_j)$ because of the properties of the matrices on the right side of (32). If the j characters $\zeta_\mu(G)$, $(\mu = \mu_1, \mu_2, \cdots, \mu_j)$, are linearly dependent for 2-regular elements G of $\mathfrak{G}$, then $M = 0$, as follows from (29).

Let us now choose a maximal system of characters of B which are linearly independent for 2-regular elements $\mathfrak{G}$. Such a system consists of y characters. We first take as many characters as possible with $\tau_\mu = m$, then as many as possible with $\tau_\mu = m + 1$, etc. We may assume that the characters chosen are $\zeta_1, \cdots, \zeta_y$. Let β_ρ be the number of characters $\zeta_1, \cdots, \zeta_y$ for which $\tau_\mu = m + \rho - 1$, $(\rho = 1, 2, \cdots, s)$, such that $\beta_1 > 0, \beta_2 \geqq 0, \cdots, \beta_{s-1} \geqq 0, \beta_s > 0$. Then

$$\begin{gathered} \tau_1 = m, \qquad \tau_y = m + s - 1; \\ m \leqq \tau_\mu \leqq m + s - 1 \quad \text{for} \quad \mu = 1, 2, \cdots, y. \end{gathered} \tag{34}$$

---

[26] BN, §9, (28).
[27] Cf. BN, (21).
[28] The determinant is actually a power of 2, cf. footnote 24.
[29] BN, §19.

Of course, we have

$$\beta_1 + \beta_2 + \cdots + \beta_s = y.$$

It now follows that every minor of degree y of $(S_{\mu\nu})$ is divisible by 2^l with

$$l = \beta_1 m + \beta_2(m+1) + \cdots + \beta_s(m+s-1) + \beta_1 + \beta_2 + \cdots + \beta_s - 1,$$

$$l = my - 1 + \sum_{\sigma=1}^{s} \sigma\beta_\sigma .$$

Similarly, the minors of degree $y-1$ are divisible by $2^{l'}$ with

$$l' = \beta_1 m + \beta_2(m+1) + \cdots + \beta_{s-1}(m+s-2) + (\beta_s - 1)(m+s-1) + \beta_1 + \cdots + \beta_s - 2,$$

$$l' = m(y-1) - 1 - s + \sum_{\sigma=1}^{s} \sigma\beta_\sigma .$$

On comparing this with (33), we find

$$ya - (\epsilon_1 + \epsilon_2 + \cdots + \epsilon_y) \geqq my - 1 + \sum_{\sigma=1}^{s} \sigma\beta_\sigma ,$$

$$(y-1)a - (\epsilon_2 + \epsilon_3 + \cdots + \epsilon_y) \geqq (y-1)m - 1 - s + \sum_{\sigma=1}^{s} \sigma\beta_\sigma .$$ [30]

From (28) we see that $r_1 r_2 \cdots r_y$ is the power of 2 with the exponent

$$(a-1-m)\beta_1 + (a-2-m)\beta_2 + \cdots + (a-s-m)\beta_s = (a-m)y - \sum_{\sigma=1}^{s} \sigma\beta_\sigma .$$

Therefore, the two inequalities can be written in the form

$$|C_\lambda| = e_1 e_2 \cdots e_y \leqq 2 r_1 r_2 \cdots r_y , \tag{35}$$

$$e_2 e_3 \cdots e_y \leqq 2^{m-a+s+1} r_1 r_2 \cdots r_y . \tag{36}$$

No two of the characters $\zeta_1, \zeta_2, \cdots, \zeta_y$ can be 2-conjugate. Letting w be the number of families of 2-conjugate characters which belong to the block $B = B_\lambda$, we have $w > y$.[31] We arrange $\zeta_{y+1}, \cdots, \zeta_x$ in such a manner that $\zeta_1, \zeta_2, \cdots, \zeta_w$ all lie in different families. Denote the μ^{th} row of D_λ by $\mathfrak{d}_\mu$, and let

$$D_\lambda = \begin{pmatrix} \mathfrak{d}_1 \\ \vdots \\ \mathfrak{d}_x \end{pmatrix} = \begin{pmatrix} \tilde{D}_\lambda \\ \mathfrak{d}_{w+1} \\ \vdots \\ \mathfrak{d}_x \end{pmatrix} = \begin{pmatrix} T \\ \mathfrak{d}_{y+1} \\ \vdots \\ \mathfrak{d}_x \end{pmatrix},$$

[30] This holds in the case $y = 1$ also; we have here $s = 1, \beta_1 = 1$, and $\epsilon_2 + \cdots + \epsilon_y$ is to be taken equal to 0, and $e_2 \cdots e_y$ equal to 1.

[31] Cf. footnote 16.

so that $\tilde{D}_\lambda$ contains the first w rows of D_λ, and T the first y rows. Then D_λ has the same rows as $\tilde{D}_\lambda$, the μ^{th} row of $\tilde{D}_\lambda$ appearing r_μ times in D_λ. Let Δ_1, Δ_2, $\cdots$ be the minors of degree y of D_λ; the determinant $|T|$ appears $r_1 r_2 \cdots r_y$ times among these minors. Consequently, we have

$$|C_\lambda| = |D_\lambda' D_\lambda| = \sum \Delta_\kappa^2 \geqq r_1 r_2 \cdots r_y |T|^2. \tag{37}$$

Here, $|T| \neq 0$, since $\zeta_1(G), \cdots, \zeta_y(G)$ are linearly independent even if G ranges over the 2-regular elements only. If all the minors of degree y of $\tilde{D}_\lambda$ except $|T|$ vanish, then $\mathfrak{d}_{y+1} = 0$, which is impossible. Hence the inequality sign must hold in (37). On comparing (35) and (37), we find

$$|T| = \pm 1, \tag{38}$$

and, since every r_ν and e_ν is a power of 2, we have

$$|C_\lambda| = 2 r_1 r_2 \cdots r_y \tag{39}$$

Then (36) yields

$$e_1 \geqq 2^{a-m-s}. \tag{40}$$

Because of (38), we may set

$$\mathfrak{d}_\nu = h_{\nu 1}\mathfrak{d}_1 + \cdots + h_{\nu y}\mathfrak{d}_y,$$

where the $h_{\nu\rho}$ are rational integers. Then $D_\lambda = HT$ with $H = (h_{\nu\rho})$. It follows that $C_\lambda = T'H'HT$, and this shows that $H'H$ has the same elementary divisors as C_λ. In particular, the coefficient in the y^{th} row, y^{th} column of $H'H$ must be divisible by e_1, and hence, by (40), we have

$$\sum_{\nu=1}^{x} h_{\nu y}^2 \equiv 0 \pmod{2^{a-m-s}}. \tag{41}$$

When ζ_ν is 2-conjugate to ζ_y, then $\mathfrak{d}_\nu = \mathfrak{d}_y$ and $h_{\nu y} = 1$. There must exist characters ζ_ν which are not 2-conjugate to ζ_y and for which $h_{\nu y} \neq 0$.[32] Using (34), we obtain

$$\sum_{\nu=1}^{x} h_{\nu y}^2 > r_y = 2^{a-1-r_y} = 2^{a-m-s}$$

and, therefore, (41) yields

$$\sum_{\nu=1}^{x} h_{\nu y}^2 \geqq 2^{a+1-m-s} = 2r_y. \tag{42}$$

Consider now the minor Δ_κ consisting of the rows $1, 2, \cdots, y-1, \nu$. Its value is $h_{\nu y}|T| = \pm h_{\nu y}$. If ζ_ν is 2-conjugate to one of the characters $\zeta_1, \cdots, \zeta_{y-1}$,

[32] In §5 of the paper mentioned in footnote 16, it is shown that it is impossible to find a linear combination of the columns of $\tilde{D}_\lambda$, such that exactly one row contains a term $\neq 0$. This implies the statement.

then the minor vanishes. It is easily seen that there are $r_1r_2 \cdots r_{y-1}$ minors of D_λ with the value $\pm h_{\nu y}$; and (37), (39) and (42) imply that

$$(43)\quad 2r_1r_2 \cdots r_y = |C_\lambda| = \sum \Delta_\kappa^2 \geqq r_1r_2 \cdots r_{y-1} \sum_{\nu=1}^{x} h_{\nu y}^2 \geqq 2r_1 \cdots r_{y-1}r_y .$$

Since actually the equality sign holds, it follows that those minors, which have not been taken into account, must all vanish. If $y > 1$, it follows that the minors formed by means of the rows $2, 3, \cdots, y, \nu$ must vanish unless ζ_ν is 2-conjugate to ζ_1. The value of such a minor is $\pm h_{\nu 1}$, and we obtain a contradiction, since there must exist characters ζ_ν, which are not 2-conjugate to ζ_1, for which $h_{\nu 1} \neq 0$.[32]

It follows that $y = 1$, i.e. that D_λ has only one column. The coefficient in the first row must be 1, because of (38). We set $h_{\nu y} = h_\nu$,

$$D_\lambda = \begin{pmatrix} 1 \\ h_2 \\ \vdots \\ h_x \end{pmatrix}.$$

The relation (43), (28), and (34) then imply

$$(44)\qquad 1 + h_2^2 + \cdots + h_x^2 = 2r_y = 2r_1 = 2^{a-m}.$$

Since ζ_1 has degree $z_1 = 2^m z'$, the degree $z_2 = 2^{\tau_2} z_2'$ of ζ_2 must be equal to $h_2 z_1$, which shows that h_2 is divisible by $2^{\tau_2 - m}$. The left side of (44) is at least equal to $r_1 + r_2 h_2^2$, since r_1 terms 1 and r_2 terms h_2 appear. Because of (28) and (34) we have

$$2^{a-m} \geqq r_1 + r_2 h_2^2 \geqq 2^{a-1-m} + 2^{a-1-\tau_2} 2^{2\tau_2 - 2m} \geqq 2^{a-m}.$$

We readily see that $w = 2$, $\tau_2 = m$, $r_1 = r_2 = 2^{a-1-m}$, $h_2 = 1$. The block B_λ consists of two families. All of its characters have the same degree $z_1 = 2^m z_1'$.[33] Since it was assumed that B contains a character of degree $2^{a-1}z'$, $((z', 2) = 1)$, it follows that $m = a - 1$ and hence $r_1 = r_2 = 1$,

$$D_\lambda = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

This proves the theorems 3, 4, 5, 6, and 7, and the corollaries of §7, for $p = 2$.

9. Applications

THEOREM 8: *If $\mathfrak{G}$ is a group of order $g = p^a q^b r^c$, (p, q, r distinct primes) with $a \leqq 2$, then $\mathfrak{G}$ possesses irreducible representations $\mathfrak{Z}$ besides the* 1*-representation* [1] *whose degree z is a power of q, and also irreducible representations $\mathfrak{Z}_1 \neq$ [1] whose degree z_1 is a power of r.*

PROOF: Consider the p-block B of representations which contains the repre-

[33] All the facts derived so far hold for any block B which contains a character ζ_ν with the following properties: (1) ζ_ν is not of the highest kind. (2) $r_\nu z_\nu \equiv 0 \pmod{2^{a-1}}$. (3) $K(\zeta_\nu)$ cannot be obtained from K by adjoining a $2^{j\text{th}}$ root of unity with $j \geqq 2$.

sentation [1]. This B is not of the highest kind, and if $a = 2$, it is not of the type $a - 1 = 1$. It follows from theorem 3 that the degrees of all representations of B are prime to p. We apply the relation[34] $\sum \zeta_\mu(R)\zeta_\mu(S) = 0$ where μ ranges over the indices belonging to B, and where R is p-regular and S is p-singular. For $R = 1$, we have $\zeta_\mu(R) = z_\mu$. The term corresponding to the character [1] is 1. At least one other term $\zeta_\mu(1)\zeta_\mu(S) = z_\mu\zeta_\mu(S)$ must be prime to r, and then z_μ must be a power of q. In the same manner we see that a character $\zeta_\nu \neq [1]$ has a degree z_ν which is a power of r.

If $\mathfrak{G}$ is equal to its commutator group $\mathfrak{G}'$, then z_μ and z_ν cannot be 1 since [1] is the only linear representation of $\mathfrak{G}$. In particular, this will be so, if G is simple.

Assume that $r^e = 4$ or $r^e = 3$. Then $\mathfrak{G}$ must have a representation of one of the degrees 2, 3, or 4. Since all linear groups of these degrees are known,[35] we obtain easily

THEOREM 9: *The only simple group of an order $4p^aq^b$, (p, q primes) with $a \leqq 2$ is the alternating group $\mathfrak{A}_5$ of order 60. The only simple groups of an order $3p^aq^b$, (p, q primes) with $a \leqq 2$ are the groups $A_5 \cong LF(2, 5)$ of order 60 and the group $LF(2, 7)$ of order 168.*

10. On $\mathfrak{P}$-similiar representations

Let $\mathfrak{Z}$ be an irreducible representation of $\mathfrak{G}$ with $\mathfrak{P}$-integral coefficients in an algebraic number field Ω, where $\mathfrak{P}$ is a prime ideal divisor of p. If $\mathfrak{Z}_1$ is a second representation with $\mathfrak{P}$-integral coefficients in the same field Ω, it may happen that $\mathfrak{Z}$ and $\mathfrak{Z}_1$ are similar, but that the corresponding modular representations $\overline{\mathfrak{Z}}$ and $\overline{\mathfrak{Z}}_1$ are not similar. Indeed, if

$$P^{-1}\mathfrak{Z}P = \mathfrak{Z}_1, \tag{45}$$

we may assume that the coefficients of P are $\mathfrak{P}$-integers which are not all divisible by $\mathfrak{P}$. Going over to residue classes mod $\mathfrak{P}$ (which again will be indicated by a bar) we obtain

$$\overline{\mathfrak{Z}}\overline{P} = \overline{P}\overline{\mathfrak{Z}}_1$$

and $\overline{P} \neq 0$, so that $\overline{\mathfrak{Z}}$ and $\overline{\mathfrak{Z}}_1$ are intertwined. However, this proves $\overline{\mathfrak{Z}} \sim \overline{\mathfrak{Z}}_1$ only if the determinant of P is not divisible by $\mathfrak{P}$. We shall say that $\mathfrak{Z}$ and $\mathfrak{Z}_1$ are $\mathfrak{P}$-similar, if P in (45) can be chosen in accordance with these conditions, i.e. with $\mathfrak{P}$-integral coefficients and a determinant $\not\equiv 0 \pmod{\mathfrak{P}}$. The class of all representations $\mathfrak{Z}_1$ which are similar to $\mathfrak{Z}$ and have $\mathfrak{P}$-integral coefficients thus breaks up into subclasses of $\mathfrak{P}$-similar representations.

Conversely, if $\mathfrak{T}$ is any representation of $\mathfrak{G}$ in the field $\bar{\Omega}$ of residue classes (mod $\mathfrak{P}$), which is similar to $\overline{\mathfrak{Z}}$, then we may find a representation $\mathfrak{Z}_1$, which is $\mathfrak{P}$-similar to $\mathfrak{Z}$, such that $\overline{\mathfrak{Z}}_1 = \mathfrak{T}$.

---

[34] R. Brauer and C. Nesbitt, University of Toronto Studies, Math. Ser. No. 4, 1937, theorem VIII.

[35] H. F. Blichfeldt, *Finite collineation groups*, Chicago 1917.

In the general case of (45), we may assume that P appears in the normal form of the theory of elementary divisors (with regard to the domain of all $\mathfrak{P}$-integers) if we replace $\mathfrak{Z}$ and $\mathfrak{Z}_1$ by $\mathfrak{P}$-similar representations. We thus may set

$$P = \begin{pmatrix} 0 & 0 & \cdots & \pi^{s-1}I_{m_s} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \pi I_{m_2} & \cdots & 0 \\ I_{m_1} & 0 & \cdots & 0 \end{pmatrix} = (\pi^{s-\varkappa}\delta_{\varkappa+\lambda,s+1}I_{m_\lambda})_{\varkappa,\lambda}, \tag{46}$$

where $m_1 > 0$, $m_2 \geqq 0, \cdots, m_{s-1} \geqq 0$, $m_s > 0$ are rational integers, I_m is the unit matrix of degree m, and π is a $\mathfrak{P}$-integer which satisfies $\pi \equiv 0 \pmod{\mathfrak{P}}$, $\pi \not\equiv 0 \pmod{\mathfrak{P}^2}$. If we set

$$\mathfrak{Z} = \begin{pmatrix} \mathfrak{A}_{11} & \mathfrak{A}_{12} & \cdots & \mathfrak{A}_{1s} \\ \mathfrak{A}_{21} & \mathfrak{A}_{22} & \cdots & \mathfrak{A}_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ \mathfrak{A}_{s1} & \mathfrak{A}_{s2} & \cdots & \mathfrak{A}_{ss} \end{pmatrix}, \qquad \mathfrak{Z}_1 = \begin{pmatrix} \mathfrak{B}_{11} & \mathfrak{B}_{12} & \cdots & \mathfrak{B}_{1s} \\ \mathfrak{B}_{21} & \mathfrak{B}_{22} & \cdots & \mathfrak{B}_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ \mathfrak{B}_{s1} & \mathfrak{B}_{s2} & \cdots & \mathfrak{B}_{ss} \end{pmatrix}, \tag{47}$$

where $\mathfrak{A}_{\varkappa\lambda}$ has $m_{s-\varkappa+1}$ rows and $m_{s-\lambda+1}$ columns, and $\mathfrak{B}_{\varkappa\lambda}$ has $m_\varkappa$ rows and m_λ columns, then (45) gives

$$\pi^{\lambda-1}\mathfrak{A}_{s+1-\varkappa,s+1-\lambda} = \pi^{\varkappa-1}\mathfrak{B}_{\varkappa\lambda}. \tag{48}$$

This shows that in (47), all the terms above the main diagonal in $\mathfrak{Z}$ and $\mathfrak{Z}_1$ are congruent to 0 (mod $\mathfrak{P}$). In the modular sense, $\mathfrak{Z}$ and $\mathfrak{Z}_1$ split into the (reducible or irreducible) constituents

$$\bar{\mathfrak{A}}_{11} = \bar{\mathfrak{B}}_{ss}, \bar{\mathfrak{A}}_{22} = \bar{\mathfrak{B}}_{s-1,s-1}, \cdots, \bar{\mathfrak{A}}_{ss} = \bar{\mathfrak{B}}_{11}.$$

If $s = 1$, then $\mathfrak{Z}$ and $\mathfrak{Z}_1$ are $\mathfrak{P}$-similar. We certainly have this case when $\mathfrak{Z}$ is modular-irreducible in $\bar{\Omega}$.

THEOREM 10: *If the representation $\mathfrak{Z}$ remains irreducible as a modular representation, then every representation $\mathfrak{Z}_1$ with $\mathfrak{P}$-integral coefficients, which is similar to $\mathfrak{Z}$, is $\mathfrak{P}$-similar to $\mathfrak{Z}$.*

Remark: This theorem can always be applied if the degree z of $\mathfrak{Z}$ is divisible by p^a.

Further, if it is known that no fixed coefficient of a representation $\mathfrak{P}$-similar to $\mathfrak{Z}$ is divisible by $\mathfrak{P}^l$ for every G in $\mathfrak{G}$, then (48) shows that $s \leqq l$.

11. On the arrangement of the modular constituents of an irreducible representation $\mathfrak{Z}$ of type $a - 1$

We shall say that the algebraic splitting field Ω of the representation $\mathfrak{Z}$ is a *normal splitting field of least ramification*, if Ω is normal over the field of rational numbers P, and if the order of ramification of the fixed rational prime p is the same for Ω as for the subfield $P(\zeta)$ obtained by adjoining the character ζ of $\mathfrak{Z}$ to P. The existence of such fields Ω follows from lemma 1. Denote by K the field of inertia of the prime ideal divisor $\mathfrak{P}$ of p. Then

$$\mathfrak{P}^r = \mathfrak{p}, \qquad (p) = \mathfrak{p}\mathfrak{q}, \quad \text{with} \quad (\mathfrak{p}, \mathfrak{q}) = (1),$$

where $\mathfrak{p}$ is a prime ideal divisor of p in K, and where r is the number of characters which are p-conjugate to ζ. As is easily seen, the degree of $K(\zeta)$ over K is r, and hence $K(\zeta) = \Omega$.

We now prove

LEMMA 8: *Let $\mathfrak{Z}$ be an irreducible representation of degree z with $z \equiv 0 \pmod{p^{a-1}}$ such that all the coefficients $a_{\kappa\lambda}(G)$ of the matrix $\mathfrak{Z}(G)$ representing G lie in the normal splitting field Ω of least ramification. For each pair (κ, λ) there exists a group element G_0 such that $a_{\kappa\lambda}(G_0)$ and $a_{\lambda\kappa}(G_0)$ are not both divisible by $\mathfrak{P}$. For each (κ, λ), there exists a group element G_0^* such that either $a_{\kappa\lambda}(G_0^*)$ is not a $\mathfrak{P}$-integer or $a_{\lambda\kappa}(G_0^*) \not\equiv 0 \pmod{\mathfrak{P}^2}$.*

PROOF: (a) Assume that $a_{\kappa\lambda}(G) \equiv a_{\lambda\kappa}(G) \equiv 0 \pmod{\mathfrak{P}}$ for every G in $\mathfrak{G}$. We now apply the method of §2 setting $\beta(G) = a_{\kappa\lambda}(G)$, $\gamma(G) = a_{\lambda\kappa}(G)$, $\omega = 1$. We have here $(r, p) = 1$ because of theorem 2, and hence $r \mid p - 1$, $m = \varphi(p)$. Then $\xi_1(G) \equiv 0$, $\xi_2(G) \equiv 0 \pmod{\mathfrak{p}}$, and (4) gives a contradiction.

(b) If $a_{\kappa\lambda}(G)$ is a $\mathfrak{P}$-integer and $a_{\lambda\kappa}(G) \equiv 0 \pmod{\mathfrak{P}^2}$, we multiply the κ^{th} row by an element $\pi \equiv 0 \pmod{\mathfrak{P}}$, for which $\pi \not\equiv 0 \pmod{\mathfrak{P}^2}$, and divide the κ^{th} column by π. The similar representation thus obtained satisfies the assumption of the part (a) of this proof. Therefore, we again obtain a contradiction.

Suppose now that the coefficients of $\mathfrak{Z}$ are $\mathfrak{P}$-integers. It follows at once from lemma 8 that for the corresponding modular representation $\overline{\mathfrak{Z}}$ all the Loewy constituents[36] are irreducible. This implies that the arrangement of the irreducible (modular) constituents of $\mathfrak{Z}$ is uniquely determined.

Further, if $\mathfrak{Z}$ and $\mathfrak{Z}_1$ both have $\mathfrak{P}$-integral coefficients and are similar, then it follows from lemma 8 that $s \leqq 2$ in the notation of §10, (46), (47). If $s = 1$, $\mathfrak{Z}$ and $\mathfrak{Z}_1$ are $\mathfrak{P}$-similar. If $s = 2$, we have by (48),

$$(49)\quad \mathfrak{Z} = \begin{pmatrix} \mathfrak{A}_{11} & \pi\mathfrak{B}_{21} \\ \mathfrak{A}_{21} & \mathfrak{A}_{22} \end{pmatrix}, \qquad P = \begin{pmatrix} 0 & \pi I_{m_2} \\ I_{m_1} & 0 \end{pmatrix}, \qquad \mathfrak{Z}_1 = \begin{pmatrix} \mathfrak{A}_{22} & \pi\mathfrak{A}_{21} \\ \mathfrak{B}_{21} & \mathfrak{A}_{11} \end{pmatrix}.$$

The modular-irreducible constituents of $\mathfrak{Z}_1$ are, of course, the same as those of $\mathfrak{Z}$, but we see that the arrangement in which they appear is a cyclic permutation H of the arrangement in $\mathfrak{Z}$; $H \neq 1$. Since all the modular constituents of $\mathfrak{Z}$ are distinct (theorem 5), it follows that $\mathfrak{Z}$ and $\mathfrak{Z}_1$ are not $\mathfrak{P}$-similar.

If the representation $\mathfrak{Z}$ with $\mathfrak{P}$-integral coefficients breaks up into two modular constituents, then we may set $\mathfrak{Z}$ in the form given by the first equation (49). If we define P, $\mathfrak{Z}_1$ by the other equations (49), then $\mathfrak{Z}_1$ is similar to $\mathfrak{Z}$. This shows that any cyclic permutation of the modular constituents of $\mathfrak{Z}$ can be effected by a transition to a similar representation $\mathfrak{Z}_1$.

THEOREM 11: *Let $\mathfrak{Z}$ be an irreducible representation of a degree $z \equiv 0 \pmod{p^{a-1}}$ whose coefficients are $\mathfrak{P}$-integers of a normal splitting field Ω of least ramification. Let j be the number of modular-irreducible constituents of $\mathfrak{Z}$. The class of all representations of $\mathfrak{G}$, whose coefficients are $\mathfrak{P}$-integers of Ω and which are similar to $\mathfrak{Z}$, splits into j subclasses of $\mathfrak{P}$-similar representations. In all representations*

[36] Cf. e.g. R. Brauer, Trans. Amer. Math. Soc. **49**, p. 502, 1941.

of a fixed subclass, the modular constituents appear in the same arrangement. The j possible arrangements are obtained from the arrangement in $\mathfrak{Z}$ by the j cyclic permutations.

It is clear that lemma 8 will, in general, not remain valid if Ω is replaced by an extension field. Then theorem 11 also will not hold.

Let Ω be the normal splitting field of least ramification mentioned in lemma 1; τ is a root of unity of an order prime to p. The splitting group of the prime ideal $\mathfrak{P}$ contains the substitution T which transforms τ into τ^p and leaves the $p^{a\,\text{th}}$ roots of unity fixed. Then T transforms any (ordinary or modular) character χ into a conjugate character χ^T. It is easily seen that if an irreducible character ζ splits into the modular characters $\varphi_\alpha, \varphi_\beta, \cdots, \varphi_\rho$ in this arrangement, then ζ^{T^κ} will split into the modular characters $\varphi_\alpha^{T^\kappa}, \varphi_\beta^{T^\kappa}, \cdots, \varphi_\rho^{T^\kappa}$. Assume again that the degree z of ζ is divisible by p^{a-1} where p is odd. If one of the constituents φ of ζ is left invariant by T^κ, then ζ and ζ^{T^κ} have a modular constituent φ in common. Since they have the same degree, they must be p-conjugate (cf. theorem 4[37]). It then follows easily that all the modular constituents of ζ will admit the substitution T^κ. The same will hold for all the characters ζ_μ which have a modular constituent in common with ζ, and finally for all the ζ_μ of the block B of ζ. Hence

THEOREM 12: *Let* B *be a block of type $a-1$, and denote by T the substitution which replaces the g'^{th} roots of unity by their p^{th} powers but leaves the $p^{a\text{th}}$ roots of unity invariant. If a power T^κ of T transforms one of the modular characters of* B *into itself, it transforms every modular character of* B *into itself; every ordinary character ζ of* B *is transformed into a p-conjugate character by T^κ (p odd).*

12. Real characters of type $a-1$

To every representation $\mathfrak{Z}$, there belongs a contragredient representation $\mathfrak{Z}^*$ of the same degree z; the characters ζ and ζ^* of $\mathfrak{Z}$ and $\mathfrak{Z}^*$ are conjugate complex, $\zeta^* = \bar{\zeta}$. If z is divisible by p^{a-1}, p odd,[38] then either ζ and ζ^* are p-conjugate, or they have no modular constituent in common, as follows from theorem 4.[37] Consequently, if ζ contains a modular character φ with $\varphi^* = \varphi$, then ζ and ζ^* must be p-conjugate. It is easily seen that, if the modular constituents of $\mathfrak{Z}$ are

$$\mathfrak{F}_1, \mathfrak{F}_2, \cdots, \mathfrak{F}_j \tag{49}$$

in this arrangement, those of $\mathfrak{Z}^*$ are

$$\mathfrak{F}_j^*, \mathfrak{F}_{j-1}^*, \cdots, \mathfrak{F}_1^*. \tag{50}$$

On the other hand, using a suitable splitting field Ω, we easily find that in any representation p-conjugate to $\mathfrak{Z}$, the modular constituents appear in the same arrangement as in $\mathfrak{Z}$. It now follows from theorem 11 that if $\zeta(G)$ is real for p-regular G the constituents (49) and (50) must be the same apart from a

[37] The number of p-conjugate characters is the same for both characters.

[38] The following theorems 13 and 14 are trivial for $p = 2$; cf. §8.

cyclic permutation. If $\mathfrak{F}_1 \sim \mathfrak{F}_\rho^*$, then $\mathfrak{F}_\nu \sim \mathfrak{F}_{\rho+1-\nu}^*$ where we set $\mathfrak{F}_{j+\mu} = \mathfrak{F}_\mu$. For odd j, there will be exactly one value of ν for which $\nu \equiv \rho + 1 - \nu \pmod{j}$, i.e. $\mathfrak{F}_\nu \sim \mathfrak{F}_\nu^*$. For even j, we have either two values ν or no value ν. Hence

THEOREM 13: *Let $\mathfrak{Z}$ be a representation of degree $z \equiv 0 \pmod{p^{a-1}}$ with the character ζ. If ζ is not real for p-regular elements, none of the modular constituents of ζ is real. If ζ is real for p-regular elements and contains j modular constituents, then, for odd j, exactly one of these modular constituents is real; if j is even, either two or none of them are real.*

If, in particular, all the modular characters of the block are real, then each ζ contains one or two modular constituents. In the tree corresponding to the block B of ζ, each vertex lies on at most two sides. Hence

THEOREM 14: *If all the modular characters of a block B of type $a - 1$ are real, then the corresponding tree is an open polygon.*

UNIVERSITY OF TORONTO
TORONTO, CANADA

ANNALS OF MATHEMATICS
Vol. 42, No. 4, October, 1941

ON THE DENSITY OF THE SUM OF SETS OF POSITIVE INTEGERS. II[1]

BY ALFRED BRAUER

In memory of I. Schur

(Received February 28, 1941)

Introduction

This paper is the continuation of my paper in the Annals of Mathematics vol. 39 (1938), pp. 322–340. Originally this continuation was to appear immediately after publication of the first part and was to contain improvements of some theorems of I. Schur by means of results of the first part. The publishing was postponed again and again, in order to improve the results. In this work I also obtained improvements of some theorems of my previous paper.

In 1, I state the results of the first part which I use in this paper.[2] In the following, let again $A_1, A_2, \cdots, A_n$ be sets of positive integers with the positive densities $\alpha_1 \leqq \alpha_2 \leqq \cdots \leqq \alpha_n$, and $\gamma < 1$ be the density of the sum $A_1 + A_2 + \cdots + A_n$. Improving some results of the first part (see below 1, Theorems V, X, and XI) I obtain instead of V the following

THEOREM: *Suppose* $\alpha_n \geqq 1 - n\alpha_1$, *then we have*

$$\gamma \geqq \frac{n-1}{n} + \frac{\alpha_n}{n}.$$

Suppose $\alpha_1 \geqq 1 - n\alpha_2$, *then we have*

$$\gamma \geqq \frac{n-1}{n} + \frac{\alpha_1}{n}.$$

If we denote $\alpha_1 + \alpha_2 + \cdots + \alpha_n$ by σ_n, we obtain instead of XI:

$$(1) \qquad \begin{cases} \gamma \geqq \frac{9}{10}\sigma_2; \quad \gamma \geqq \frac{81}{92}\sigma_3 > .8804\sigma_3; \quad \gamma \geqq \frac{9712}{11117}\sigma_4 > .8701\sigma_4; \\ \gamma \geqq \frac{340220}{391897}\sigma_n > .8679\sigma_n \quad \text{for } n \geqq 5. \end{cases}$$

Moreover I obtain instead of X:

$$\gamma \geqq \alpha_1 + \tfrac{5}{6}\alpha_2.$$

If the conjecture is correct that always

$$(2) \qquad \gamma \geqq \alpha_1 + \alpha_2 + \cdots + \alpha_n,$$

[1] Presented to the American Mathematical Society, December 27, 1939.

[2] An interesting survey on the problems connected with this paper was given by H. Rohrbach in his paper *Einige neuere Untersuchungen über die Dichte in der additiven Zahlentheorie*, Jahresbericht der Deutschen Mathematiker-Vereinigung v. 48 (1938), pp. 199–236.

then one would obtain immediately for every n and m with $n \geqq m$ that

$$\gamma \geqq \frac{n}{m}(\alpha_1 + \alpha_2 + \cdots + \alpha_m). \tag{3}$$

In particular, it follows from (2) that the sum of the densities satisfies the inequality

$$\alpha_1 + \alpha_2 + \cdots + \alpha_n < 1 \tag{4}$$

for $\gamma < 1$, and more generally

$$\alpha_1 + \alpha_2 + \cdots + \alpha_m < m/n. \tag{5}$$

Even the special cases (4) and (5) are still unproved. One only knows that (4) is true for $n = 2$ and (5) moreover for $m = 1$. The former follows from the theorem of Schnirelmann, the latter from the theorem of Khintchine (see below, Theorems I and II).

In the following, I prove that (5) is true at all events if n and m satisfy the condition

$$\frac{1}{n-1} + \frac{1}{n-2} + \cdots + \frac{1}{n-m+1} \leqq \frac{m}{n}. \tag{6}$$

For $m = 2$, the condition (6) is always satisfied; for $m = 3$ it is satisfied if $n \geqq 5$, and for fixed m certainly if

$$n \geqq \tfrac{1}{2}(m^2 + m - 2).$$

Moreover I prove that (5) holds also for $m = 3$ and $n = 4$.

It follows from the theorem of Khintchine that the sum of n sets $A_1 + A_2 + \cdots + A_n$ is the set of all positive integers if $\alpha_1 \geqq 1/n$. It now follows that the same is true if

$$\alpha_1 + \alpha_2 + \cdots + \alpha_m \geqq m/n$$

where m and n satisfy the condition (6), or where $m = 3$, $n = 4$.

I. Schur[3] proved

$$\sigma_3 < 1.3191; \quad \sigma_4 < 1.3913; \quad \sigma_5 < 1.4360;$$

$$\sigma_n < 1.6263 \quad \text{for } n > 5,$$

in approximation to the conjecture (4). I improve this result as follows:

$$\sigma_3 < 21/20 = 1.05; \quad \sigma_4 < 97/90 < 1.0778; \quad \sigma_5 < 43/39 < 1.1026;$$

$$\sigma_n < \frac{39197n}{34020n + 5177} < 1.1522 \quad \text{for } n > 5.$$

[3] I. Schur, Über den Begriff der Dichte in der additiven Zahlentheorie, Sitzungsberichte der Preussischen Akademie der Wissenschaften Phys.-Math. Klasse 1936, pp. 269–297.

In order to approximate (2) and (3) for every n and m with $m \leqq n$ we shall determine constants $c_n^{(m)}$ as large as possible such that

$$\gamma \geqq c_n^{(m)}(\alpha_1 + \alpha_2 + \cdots + \alpha_m).$$

For the special case $n = m$ we obtain such constants from (1). I. Schur[4] proved that one can choose

$$(7) \qquad c_n^{(1)} = n \quad \text{and} \quad c_n^{(m)} = \frac{nc_{n-1}^{(m-1)}}{c_{n-1}^{(m-1)} + (n^2 - n)^{\frac{1}{2}}} \quad \text{for } m > 1.$$

In the following I shall show that we can choose $c_n^{(1)} = n$ and for $m > 1$

$$(8) \qquad \begin{cases} c_n^{(m)} = \dfrac{n(n+2)c_{n-1}^{(m-1)}}{(n+2)c_{n-1}^{(m-1)} + n^2 + n - 1} & \text{if } c_{n-1}^{(m-1)} > \dfrac{n+1}{n+2}, \\ c_n^{(m)} = c_{n-1}^{(m-1)} & \text{if } c_{n-1}^{(m-1)} \leqq \dfrac{n+1}{n+2}. \end{cases}$$

It will be proved that this gives a better estimate than (7) for $m \geqq 2$.

For $n > m$ we set $n - m = d$; for a fixed value of d we denote by z the greatest integer such that

$$\log(z-1) + 1/(z+1) \leqq d + \log d + 2.$$

It then follows from (8) that

$$(9) \qquad \begin{cases} c_n^{(m)} > \dfrac{n+1}{m + \log n - \log d} & \text{for } n \leqq z, \\ c_n^{(m)} > \dfrac{z+1}{z+2} > \dfrac{e^{d+1}d + 1}{e^{d+1}d + 2} & \text{for } n \geqq z. \end{cases}$$

in approximation to (3).

Let m be fixed and n sufficiently large. I. Schur proved

$$c_n^{(m)} > n/m - 1/4$$

whereas I obtain the better estimate

$$c_n^{(m)} > n/m - \epsilon$$

from (9) for every $\epsilon > 0$.

1. Summary of former results

DEFINITION 1: *Let $A_1, A_2, \cdots, A_n$ be sets of positive integers. The sum $S = A_1 + A_2 + \cdots + A_n$ is the set of all different positive integers representable in the form*

$$e_1a_1 + e_2a_2 + \cdots + e_na_n \quad \text{with} \quad a_\nu \text{ in } A_\nu; \qquad e_\nu = 0, 1 \quad \text{for} \quad \nu = 1, 2, \cdots, n.$$

[4] L. c., footnote 2.

DEFINITION 2: *Let $A(x)$ be the number of elements of A less than x. The density α of A is the lower bound of $A(x)/x$ for $x = 1, 2, \cdots$.*

THEOREM I (Schnirelmann): *If $n = 2$ and $\alpha_1 + \alpha_2 \geqq 1$, then we have $\gamma = 1$.*

In the following we assume $\gamma < 1$.

THEOREM II (Khintchine): $\gamma \geqq n\alpha_1$.

THEOREM III (Khintchine): *Suppose $0 \leqq \mu < \alpha < 1/n$ and for $x = 1, 2, \cdots$*

$$A_\nu(x) \geqq \alpha x - \frac{\nu - 1}{n} \qquad (\nu = 1, 2, \cdots n-1),$$

$$A_n(x) \geqq (1 - n\alpha)x - n\mu.$$

Denote by S the sum $A_1 + A_2 + \cdots + A_n$, then

$$S(x) \geqq (1 - \alpha)x - \mu.$$

THEOREM IV (Schur): $\gamma \geqq \alpha_2/(1 - \alpha_1) \qquad (n = 2).$

THEOREM V [1][5]: *If $\alpha_1 < 1/n$ and $\alpha_n \geqq 1 - n\alpha_1$, then we have*

$$\gamma \geqq 1 - \alpha_1.$$

If $\alpha_2 < 1/n$ and $\alpha_1 \geqq 1 - n\alpha_2$, then we have

$$\gamma \geqq 1 - \alpha_2.$$

THEOREM VI [6]: $$\gamma \geqq \alpha_\nu + \sum_{\substack{\mu=1 \\ \mu \neq \nu}}^{n} \frac{1}{1 + [1/\alpha_\mu]}.$$

THEOREM VII [14]: *Let $k \geqq 3$ be any integer and $k/(k-1)^2 \geqq \alpha_1$ then we have for $n = 2$*

$$\gamma \geqq \alpha_2 + \frac{k-1}{k}\alpha_1.$$

THEOREM VIII [15] and [16]: *Denote by β the density of $A_2 + A_3 + \cdots + A_n$. Then we have*

$$\gamma \geqq \beta + \frac{n+1}{n+2}\alpha_1 \geqq \alpha_n + \frac{3}{4}\alpha_{n-1} + \frac{4}{5}\alpha_{n-2} + \cdots + \frac{n+1}{n+2}\alpha_1.$$

THEOREM IX [17]: *For every a with $0 < a < 1$ we have*

$$\gamma \geqq a\beta + \alpha_1\left(n - \frac{n^2 + n - 1}{n+2}a\right).$$

THEOREM X [17a]: $\gamma \geqq \alpha_1 + \frac{4}{5}\alpha_2.$

THEOREM XI [18]: $$\gamma \geqq \frac{8}{9}(\alpha_1 + \alpha_2); \qquad \gamma \geqq \frac{120}{139}(\alpha_1 + \alpha_2 + \alpha_3);$$

$$\gamma \geqq \frac{2880}{3361}(\alpha_1 + \alpha_2 + \cdots + \alpha_n) \quad \text{for } n \geqq 4.$$

[5] The number in parentheses refers to the numeration in the first part of this paper, Annals of Mathematics v. 39 (1938), pp. 322–340.

In what follows these theorems will be cited by I, II, and so on.

2. Proof of the conjecture (5) for small values of m

First we prove

THEOREM 1: *Let n and m satisfy the conditions $1 < m \leqq n$ and*

$$\frac{1}{n-1} + \frac{1}{n-2} + \cdots + \frac{1}{n-m+1} \leqq \frac{m}{n}. \tag{10}$$

Then we have for $\gamma < 1$

$$\alpha_1 + \alpha_2 + \cdots + \alpha_m < \frac{m}{n}. \tag{11}$$

PROOF: For $m = n$ the condition (10) is satisfied only for $m = n = 2$. In this case, (11) follows from I. Therefore we may now assume that $m < n$.

Let r be any integer with $0 \leqq r \leqq m$. Denote by β_r the density of $A_1 + A_2 + \cdots + A_r$ and by β_r^* the density of $A_{r+1} + A_{r+2} + \cdots + A_n$; we set $\beta_0 = 0$. Then it follows from II that

$$\beta_r \geqq r\alpha_1 \quad \text{and} \quad \beta_r^* \geqq (n-r)\alpha_{r+1}. \tag{12}$$

Because of I and $\gamma < 1$ we have $\beta_r + \beta_r^* < 1$, and it follows from (12) that

$$r\alpha_1 + (n-r)\alpha_{r+1} < 1,$$

$$\frac{r\alpha_1}{n-r} + \alpha_{r+1} < \frac{1}{n-r}. \tag{13}$$

We add the inequalities (13) for $r = 0, 1, \cdots, m-1$ after multiplying the first of them by a non-negative number w which will be determined later. Then we obtain

$$\alpha_1\left(w + \frac{1}{n-1} + \frac{2}{n-2} + \cdots + \frac{m-1}{n-m+1}\right) + \alpha_2 + \alpha_3 + \cdots + \alpha_m < \frac{w}{n} + \frac{1}{n-1} + \frac{1}{n-2} + \cdots + \frac{1}{n-m+1}. \tag{14}$$

On the other hand, it follows from (10) that

$$\frac{n}{n-1} + \frac{n}{n-2} + \cdots + \frac{n}{n-m+1} \leqq m,$$

$$\frac{n}{n-1} - 1 + \frac{n}{n-2} - 1 + - \cdots + \frac{n}{n-m+1} - 1 \leqq 1,$$

$$\frac{1}{n-1} + \frac{2}{n-2} + \cdots + \frac{m-1}{n-m+1} \leqq 1. \tag{15}$$

Therefore, setting

$$w = 1 - \frac{1}{n-1} - \frac{2}{n-2} - \cdots - \frac{m-1}{n-m+1}$$

we have $w \geqq 0$, hence

$$\alpha_1 + \alpha_2 + \cdots + \alpha_m < \frac{1}{n}\left(1 - \frac{1}{n-1} - \frac{2}{n-2} \cdots - \frac{m-1}{n-m+1}\right) + \frac{1}{n-1} + \frac{1}{n-2} + \cdots + \frac{1}{n-m+1} = \frac{m}{n}$$

because of (14), and the theorem is proved.

It is easy to examine the condition (10) for small values of m. For $m = 2$ it is always satisfied, for $m = 3$ if $n \geqq 5$, and so on. Since (10) is equivalent to (15) and since

$$\frac{1}{n-1} + \frac{2}{n-2} + \cdots + \frac{m-1}{n-m+1} < \frac{1+2+\cdots+m-1}{n-m+1} = \frac{m(m-1)}{2(n-m+1)}$$

the inequality (11) is proved if

$$\frac{m(m-1)}{2(n-m+1)} \leqq 1,$$

$$n \geqq \tfrac{1}{2}m(m-1) + m - 1 = \tfrac{1}{2}(m^2 + m - 2).$$

COROLLARY 1: *If* (10) *holds and* (11) *is not satisfied, then the sum is the set of all positive integers.*

COROLLARY 2: *Let m, k and t be integers, $m \leqq t$ and $n = kt$; if*

$$\frac{1}{t-1} + \frac{1}{t-2} + \cdots + \frac{1}{t-m+1} \leqq \frac{m}{t},$$

then we have

$$\alpha_1 + \alpha_{k+1} + \alpha_{2k+1} + \cdots + \alpha_{(m-1)k+1} < \frac{m}{n}.$$

PROOF: Let $B_{\mu+1}$ be the sum $A_{\mu k+1} + A_{\mu k+2} + \cdots + A_{(\mu+1)k}$ for $\mu = 0, 1, \cdots, t-1$ and $\beta_{\mu+1}$ the density of $B_{\mu+1}$. It follows from II that

$$\beta_{\mu+1} \geqq k\alpha_{\mu k+1}. \tag{16}$$

Applying Theorem 1 to the sets $B_1, B_2, \cdots, B_t$ we have

$$\beta_1 + \beta_2 + \cdots + \beta_m < \frac{m}{t},$$

therefore, because of (16),

$$\alpha_1 + \alpha_{k+1} + \cdots + \alpha_{(m-1)k+1} < \frac{m}{kt} = \frac{m}{n}.$$

In particular, it follows for $m = 2$ and $n = 2k$ that

$$\alpha_1 + \alpha_{k+1} < \frac{2}{n}.$$

For odd n we infer

COROLLARY 3: *For $n = 2k + 1$ and $m = 2$ we have*

$$\alpha_1 + \alpha_{k+1} < \frac{2}{n}.$$

PROOF: We have

$$k\alpha_1 + (k + 1)\alpha_{k+1} < 1,$$

hence

$$(k + \tfrac{1}{2})(\alpha_1 + \alpha_{k+1}) < 1$$

because of $\alpha_1 \leqq \alpha_{k+1}$. This proves Corollary 3.

COROLLARY 4: *For $m = 3$ and $n = 4$ we have*

$$\alpha_1 + \alpha_2 + \alpha_3 < \tfrac{3}{4}.$$

PROOF: Suppose $\alpha_1 + \alpha_2 + \alpha_3 \geqq \frac{3}{4}$. Then we must have $\alpha_2 > \frac{1}{4}$ since $\alpha_1 + \alpha_3 < \frac{1}{2}$ by Corollary 2. It now follows from VI that the density α_{12} of $A_1 + A_2$ satisfies the condition

$$\alpha_{12} \geqq \alpha_1 + \tfrac{1}{4}.$$

Therefore we have by I and II

$$\alpha_{12} + 2\alpha_3 < 1,$$

$$\alpha_1 + 2\alpha_3 < 1 - \tfrac{1}{4},$$

$$\alpha_1 + \alpha_2 + \alpha_3 < \tfrac{3}{4}$$

since $\alpha_2 \leqq \alpha_3$. This gives a contradiction, and the corollary is proved.

3. Generalisation of Theorem VIII

THEOREM 2: *Let $\alpha_1 \leqq \alpha_2 \leqq \cdots \leqq \alpha_n$ be the densities of the sets $A_1, A_2, \cdots, A_n$ and $\gamma < 1$ be the density of $A_1 + A_2 + \cdots + A_n$. Let $\alpha_{k-l,k}$ be the density of $A_{k-l} + A_{k-l+1} + \cdots + A_k$ for $n \geqq k > l \geqq 1$. Then we have*

$$\alpha_{k-l,k} \geqq \alpha_k + \frac{n-k+3}{n-k+4}\alpha_{k-1} + \frac{n-k+4}{n-k+5}\alpha_{k-2} + \cdots + \frac{n-k+l+2}{n-k+l+3}\alpha_{k-l}.$$

For $k = n$ and $l = n - 1$ we obtain Theorem VIII.

PROOF: In the first place we consider the case $l = 1$; we have to prove

$$\alpha_{k-1,k} \geqq \alpha_k + \frac{n-k+3}{n-k+4}\alpha_{k-1}.$$

If

$$\alpha_{k-1} \leqq (n-k+4)/(n-k+3)^2,$$

the statement follows from VII. Suppose now

$$\alpha_{k-1} > (n-k+4)/(n-k+3)^2.$$

Then we have

$$\frac{(n-k+3)^2}{n-k+4}\,\alpha_{k-1} > 1, \tag{17}$$

It follows from I and II that

$$(n-k+1)\alpha_k + \alpha_{k-1} < 1. \tag{18}$$

By (17) and (18) we have

$$\begin{aligned}
\alpha_k + \frac{n-k+3}{n-k+4}\,\alpha_{k-1} &< \frac{1}{n-k+1} - \frac{\alpha_{k-1}}{n-k+1} + \frac{n-k+3}{n-k+4}\,\alpha_{k-1} \\
&< \frac{(n-k+3)^2}{(n-k+1)(n-k+4)}\,\alpha_{k-1} - \frac{\alpha_{k-1}}{n-k+1} + \frac{(n-k+1)(n-k+3)}{(n-k+1)(n-k+4)}\,\alpha_{k-1} \\
&= \frac{(n-k+3)(2n-2k+4)}{(n-k+1)(n-k+4)}\,\alpha_{k-1} - \frac{\alpha_{k-1}}{n-k+1} \\
&= 2\alpha_{k-1}\left\{1 + \frac{2}{(n-k+1)(n-k+4)}\right\} - \frac{\alpha_{k-1}}{n-k+1} \leqq 2\alpha_{k-1} \leqq \alpha_{k-1,k}
\end{aligned}$$

because of II.

Let us assume now that the theorem is proved for every k and for $l = 1, 2, \cdots, \lambda - 1$ with $\lambda < k$. Then we have

$$\alpha_{k-\lambda+1,k} \geqq \alpha_k + \frac{n-k+3}{n-k+4}\,\alpha_{k-1} + \cdots + \frac{n-k+\lambda+1}{n-k+\lambda+2}\,\alpha_{k-\lambda+1} \tag{19}$$

We have to prove the theorem for $l = \lambda$. If

$$\alpha_{k-\lambda} \leqq (n-k+\lambda+3)/(n-k+\lambda+2)^2,$$

the statement follows again from VII and (19). Therefore let us suppose

$$\alpha_{k-\lambda} > (n-k+\lambda+3)/(n-k+\lambda+2)^2. \tag{20}$$

By I and II we have

$$(n-k+1)\alpha_k + \lambda\alpha_{k-\lambda} < 1. \tag{21}$$

First we shall prove the relation

$$\alpha_k + \frac{n-k+\lambda+2}{n-k+\lambda+3}\,\alpha_{k-\lambda} < 2\alpha_{k-\lambda}. \tag{22}$$

As in the case where $\lambda = 1$ we obtain from (20) and (21)

$$\alpha_k + \frac{n-k+\lambda+2}{n-k+\lambda+3}\alpha_{k-\lambda} < \frac{1}{n-k+1} - \frac{\lambda}{n-k+1}\alpha_{k-\lambda} + \frac{n-k+\lambda+2}{n-k+\lambda+3}\alpha_{k-\lambda}$$

$$< \alpha_{k-\lambda}\frac{(n-k+\lambda+2)(2n-2k+\lambda+3)-\lambda(n-k+\lambda+3)}{(n-k+1)(n-k+\lambda+3)}$$

$$= \frac{2n^2-4nk+2n\lambda+7n+2k^2-2k\lambda-7k+2\lambda+6}{n^2-2nk+n\lambda+4n+k^2-k\lambda-4k+\lambda+3}\alpha_{k-\lambda} \leqq 2\alpha_{k-\lambda}$$

since $n \geqq k$.

Next, we state that under the assumption (20)

$$\frac{n-k+\mu+2}{n-k+\mu+3}\alpha_{k-\mu} < \alpha_{k-\lambda} \qquad \text{for } \mu = 1, 2, \cdots, \lambda-1. \tag{23}$$

As in (21) we have

$$(n-k+\mu+1)\alpha_{k-\mu} + (\lambda-\mu)\alpha_{k-\lambda} < 1.$$

Therefore, because of (20),

$$\begin{cases} \alpha_{k-\mu} < \dfrac{1-(\lambda-\mu)\alpha_{k-\lambda}}{n-k+\mu+1} < \dfrac{(n-k+\lambda+2)^2-(\lambda-\mu)(n-k+\lambda+3)}{(n-k+\mu+1)(n-k+\lambda+3)}\alpha_{k-\lambda} \\ \quad = \dfrac{n^2-2nk+n\lambda+4n+k^2-k\lambda-4k+\lambda+4+\mu n-\mu k+\mu\lambda+3\mu}{n^2-2nk+n\lambda+4n+k^2-k\lambda-4k+\lambda+3+\mu n-\mu k+\mu\lambda+3\mu}\alpha_{k-\lambda}. \end{cases} \tag{24}$$

Putting

$$n-k+\mu+3 = r$$

and

$$n^2-2nk+n\lambda+4n+k^2-k\lambda-4k+\lambda+3+\mu n-\mu k+\mu\lambda+3\mu = s$$

we obtain from (24)

$$\frac{n-k+\mu+2}{n-k+\mu+3}\alpha_{k-\mu}$$

$$= \frac{r-1}{r}\alpha_{k-\mu} < \frac{(r-1)(s+1)}{rs}\alpha_{k-\lambda} = \left(1+\frac{r-s-1}{rs}\right)\alpha_{k-\lambda} < \alpha_{k-\lambda}$$

if

$$r < s+1.$$

But this is true, since the following three inequalities hold

$$n-k \leqq n^2-2nk+k^2,$$

$$k(4+\lambda+\mu) \leqq n(4+\lambda+\mu),$$

$$\mu+3 < 3\mu+4+\mu\lambda+\lambda.$$

Therefore (23) is proved.

Finally it follows from (22) and (23) that

$$\alpha_k + \frac{n-k+3}{n-k+4}\alpha_{k-1} + \cdots + \frac{n-k+\lambda+2}{n-k+\lambda+3}\alpha_{k-\lambda} < 2\alpha_{k-\lambda} + (\lambda-1)\alpha_{k-\lambda}$$

$$= (\lambda+1)\alpha_{k-\lambda} \leqq \alpha_{k-\lambda,k}$$

because of II. Hence Theorem 2 is proved for $l = \lambda$ and consequently for each l.

4. Some lemmas

LEMMA 1: *Suppose* $n = 2$ *and*

$$\alpha_1 + \alpha_2 < 2/3. \tag{25}$$

We put

$$\gamma^* = \operatorname{Max}\left\{2\alpha_1,\ \alpha_2 + \frac{1}{1+[1/\alpha_1]},\ \alpha_1 + \frac{1}{1+[1/\alpha_2]}\right\}; \tag{26}$$

then we have

$$\alpha_1 + \alpha_2 - \gamma^* \leqq .05. \tag{27}$$

I.e., the difference between $\alpha_1 + \alpha_2$ and the lower bound for the density of $A_1 + A_2$ obtained by II and VI is less than .05, if (25) is satisfied.

PROOF: If (27) is not satisfied, we have

$$\operatorname{Min}\left\{\alpha_2 - \alpha_1,\ \alpha_1 - \frac{1}{1+[1/\alpha_1]},\ \alpha_2 - \frac{1}{1+[1/\alpha_2]}\right\} > .05 \tag{28}$$

because of (26). But the difference $1/k - 1/(k+1)$, where $k \geqq 4$ is an integer, is less than or equal to $\frac{1}{4} - \frac{1}{5} = \frac{1}{20}$. Hence

$$\alpha_1 - \frac{1}{1+[1/\alpha_1]} \leqq \frac{1}{[1/\alpha_1]} - \frac{1}{1+[1/\alpha_1]} \leqq \frac{1}{20} \qquad \text{for } \alpha_1 \leqq \tfrac{1}{4}. \tag{29}$$

Moreover (29) is true for $\frac{1}{4} < \alpha_1 \leqq \frac{1}{4} + \frac{1}{20} = .3$; therefore we have $\alpha_1 > .3$ and $\alpha_2 > .3 + .05 > \frac{1}{3}$ because of (28). Thus $\alpha_2 > \frac{1}{3} + .05$ again because of (28), and we obtain

$$\alpha_1 + \alpha_2 > .3 + \tfrac{1}{3} + .05 > \tfrac{2}{3}.$$

This contradicts (25). Therefore (27) is valid.

LEMMA 2: *Suppose* $n = 4$ *and* $\gamma < 1$. *Denote by* γ^*_{123} *the lower bound for the density of* $A_1 + A_2 + A_3$ *obtained by the theorems of this paper. Then we have*

$$d = \alpha_1 + \alpha_2 + \alpha_3 - \gamma^*_{123} \leqq \tfrac{7}{30}.$$

PROOF: We write

$$\delta_\nu = \alpha_\nu - \frac{1}{1+[1/\alpha_\nu]} \qquad \text{for } \nu = 1, 2, 3.$$

Then by VI

$$d \leqq \operatorname*{Min}_{\kappa \neq \lambda} (\delta_\kappa + \delta_\lambda). \tag{30}$$

It follows from Theorem 2 that the density α_{12} of $A_1 + A_2$ satisfies the inequality

$$\alpha_{12} \geqq \alpha_2 + \tfrac{5}{6}\alpha_1 . \tag{31}$$

On the other hand we have by I, II, and $\gamma < 1$

$$\alpha_{12} + 2\alpha_3 < 1, \tag{32}$$

hence by (31)

$$2\alpha_3 + \alpha_2 + \tfrac{5}{6}\alpha_1 < 1. \tag{33}$$

We have now to distinguish between some cases.

1) $\qquad \alpha_1 \geqq \frac{121}{540}.$

Because of Corollary 4 of Theorem 1 we have $\alpha_1 + \alpha_2 + \alpha_3 < \frac{3}{4}$ and by II

$$d \leqq \alpha_1 + \alpha_2 + \alpha_3 - 3\alpha_1 < \frac{135 - 121}{180} = \frac{7}{90}.$$

Therefore we may assume

$$\alpha_1 < \tfrac{121}{540} < \tfrac{7}{30},$$

$$\delta_1 \leqq \tfrac{1}{30}. \tag{34}$$

2) $\qquad \alpha_2 \leqq \frac{1}{5} + \frac{2}{45} = \frac{11}{45}.$

We have by (30) and (34)

$$d \leqq \delta_2 + \delta_1 \leqq \tfrac{2}{45} + \tfrac{1}{30} = \tfrac{7}{90}.$$

Hence we may assume

$$\alpha_2 > \tfrac{11}{45} \tag{35}$$

and because of (33)

$$\alpha_3 < (1 - \alpha_2)/2 = \tfrac{17}{45}. \tag{36}$$

3) $\qquad \alpha_3 > \frac{1}{3}.$

It follows from (33) and (35) that

$$\alpha_1 < \tfrac{6}{5}(1 - 2\alpha_3 - \alpha_2) < \tfrac{6}{5}(1 - \tfrac{2}{3} - \tfrac{11}{45}) = \tfrac{8}{75} < \tfrac{1}{9}.$$

Therefore we have $\delta_1 < \frac{1}{90}$ and $\delta_3 < \frac{17}{45} - \frac{1}{3} = \frac{2}{45}$ because of (36), hence $d < \frac{7}{90}$ because of (30). Therefore we may assume

$$\alpha_3 \leqq \tfrac{1}{3}. \tag{37}$$

4) $$\alpha_2 \geqq \tfrac{13}{45} \quad \text{or} \quad \alpha_3 \leqq \tfrac{13}{45}.$$

We have by II, (37), (35), and (34)

$$d \leqq \alpha_3 - \alpha_2 + \delta_1 \leqq \tfrac{2}{45} + \tfrac{1}{30} = \tfrac{7}{90}.$$

Hence we may assume

(38) $$\alpha_2 < \tfrac{13}{45},$$

(39) $$\alpha_3 > \tfrac{13}{45} > \tfrac{1}{4}.$$

5) $$\alpha_3 \leqq \tfrac{1}{4} + \tfrac{2}{45}.$$

We have by (34) and (39)

$$d \leqq \delta_1 + \delta_3 \leqq \tfrac{1}{30} + \tfrac{2}{45} = \tfrac{7}{90}.$$

Hence we may assume

(40) $$\alpha_3 > \tfrac{53}{180}.$$

6) $$\alpha_1 \leqq \tfrac{7}{36}.$$

We have $\delta_1 \leqq \frac{1}{36}$ and $\delta_2 \leqq \frac{1}{20}$ because of (38). Therefore we may assume

$$\alpha_1 > \tfrac{7}{36}$$

and obtain by VI

$$\alpha_{12} \geqq \alpha_2 + \tfrac{1}{6},$$

$$\alpha_{12} + 2\alpha_3 > \tfrac{1}{6} + \tfrac{11}{45} + \tfrac{53}{90} = 1$$

because of (35) and (40). This contradicts (32) and Lemma 2 is proved.

5. Estimates for the sum of the densities in the cases $n = 3$ and $n = 4$

THEOREM 3: *Let $n = 3$. If $A_1 + A_2 + A_3$ is not the set of all positive integers, then we have*

(41) $$\alpha_1 + \alpha_2 + \alpha_3 < 1.05.$$

PROOF: Let γ_{12}^* be the lower bound for the density of $A_1 + A_2$ obtained by II and VI. Because of Theorem 1 we have $\alpha_1 + \alpha_2 < \frac{2}{3}$. Therefore it follows from Lemma 1 that

(42) $$\alpha_1 + \alpha_2 - \gamma_{12}^* < .05.$$

On the other hand we have by I

(43) $$\gamma_{12}^* + \alpha_3 < 1.$$

Adding (42) and (43) we obtain (41).

THEOREM 4: *Let $n = 4$ be. If $A_1 + A_2 + A_3 + A_4$ is not the set of all positive integers, then we have*

(44) $$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 < \tfrac{97}{90} < 1.0778.$$

PROOF: Let γ^*_{123} be the lower bound for the density of $A_1 + A_2 + A_3$ obtained by the theorems of this paper. It follows from Lemma 2 that

$$\alpha_1 + \alpha_2 + \alpha_3 - \gamma^*_{123} < \tfrac{7}{90} \tag{45}$$

and from I that

$$\gamma^*_{123} + \alpha_4 < 1. \tag{46}$$

By the addition of (45) and (46) we obtain (44).

6. Improvement of the Theorems V, X, and XI

THEOREM 5: *If the densities* $\alpha_1 \leqq \alpha_2 \leqq \cdots \leqq \alpha_n$ *of* $A_1, A_2, \cdots, A_n$ *satisfy the condition*

$$\alpha_n \geqq 1 - n\alpha_1, \tag{47}$$

then we have

$$\gamma \geqq \frac{n-1}{n} + \frac{\alpha_n}{n};$$

if

$$\alpha_1 \geqq 1 - n\alpha_2, \tag{48}$$

then

$$\gamma \geqq \frac{n-1}{n} + \frac{\alpha_1}{n}.$$

PROOF: If (47) is satisfied, we set

$$\alpha = (1 - \alpha_n)/n, \tag{49}$$

then $\alpha < 1/n$ and by (47)

$$\alpha_\nu \geqq \alpha_1 \geqq (1 - \alpha_n)/n = \alpha \quad (\nu = 2, 3, \cdots, n-1),$$

$$\alpha_n = 1 - n\alpha.$$

It follows that for $x = 1, 2, \cdots$

$$A_\nu(x) \geqq \alpha_\nu x \geqq \alpha x \geqq \alpha x - (\nu - 1)/n \qquad (\nu = 1, 2, \cdots, n-1),$$

$$A_n(x) \geqq (1 - n\alpha)x.$$

Applying III for $\mu = 0$ to the sets $A_1, A_2, \cdots, A_n$ we obtain by (49)

$$\gamma \geqq 1 - \alpha = 1 - \frac{1 - \alpha_n}{n} = \frac{n-1}{n} + \frac{\alpha_n}{n}.$$

Suppose now that (48) is satisfied. Here we set

$$\alpha = (1 - \alpha_1)/n. \tag{50}$$

Then we have $\alpha < 1/n$ and

$$\alpha_\nu \geqq \alpha_2 \geqq (1 - \alpha_1)/n = \alpha \qquad (\nu = 3, 4, \cdots, n),$$

$$\alpha_1 = 1 - n\alpha.$$

Applying III to the sets $A_2, A_3, \cdots, A_n, A_1$ we obtain by (50)

$$\gamma \geqq 1 - \alpha = 1 - \frac{1 - \alpha_1}{n} = \frac{n-1}{n} + \frac{\alpha_1}{n}.$$

By means of Theorem 5 it is possible to improve XI. First we consider the case $n = 2$.

THEOREM 6: *For $n = 2$ we have $\gamma \geqq \frac{9}{10}(\alpha_1 + \alpha_2)$.*

PROOF: We have to distinguish between certain cases:

1) $$\alpha_1 \leqq \tfrac{5}{16}.$$

It follows from VII that

$$\gamma \geqq \alpha_2 + \tfrac{4}{5}\alpha_1$$

and from II that

$$\gamma \geqq 2\alpha_1,$$

hence

$$10\gamma + \gamma \geqq 10(\alpha_1 + \alpha_2),$$

$$\gamma \geqq \tfrac{10}{11}(\alpha_1 + \alpha_2) > \tfrac{9}{10}(\alpha_1 + \alpha_2).$$

Therefore we may assume that

(51) $$\tfrac{1}{2} > \alpha_1 > \tfrac{5}{16}.$$

2) $$\alpha_1 \geqq \tfrac{9}{11}\alpha_2.$$

We deduce from II

$$\gamma \geqq 2\alpha_1 = \tfrac{9}{10}\alpha_1 + \tfrac{11}{10}\alpha_1 \geqq \tfrac{9}{10}(\alpha_1 + \alpha_2).$$

Hence we may assume that

(52) $$\alpha_1 < \tfrac{9}{11}\alpha_2$$

and, because of (51)

(53) $$\alpha_2 > \tfrac{11}{9} \cdot \tfrac{5}{16} = \tfrac{55}{144}.$$

3) $$9\alpha_1 + 4\alpha_2 \leqq 5.$$

It follows from (51) and (53) that

$$2\alpha_1 + \alpha_2 \geqq \tfrac{145}{144} > 1,$$

hence we have because of Theorem 5 for $n = 2$

$$\gamma \geqq \tfrac{1}{2} + \frac{\alpha_2}{2} \geqq \tfrac{9}{10}\alpha_1 + \tfrac{4}{10}\alpha_2 + \tfrac{1}{2}\alpha_2 = \tfrac{9}{10}(\alpha_1 + \alpha_2).$$

Therefore we may assume

(54) $$9\alpha_1 + 4\alpha_2 > 5.$$

4) $$\alpha_1 \leqq \tfrac{1}{3}.$$

By VI and (51) we have

(55) $$\gamma \geqq \tfrac{1}{4} + \alpha_2 = \tfrac{1}{4} + \tfrac{1}{10}\alpha_2 + \tfrac{9}{10}\alpha_2 .$$

On the other hand, it follows from (54) that

$$\tfrac{1}{10}\alpha_2 > \tfrac{1}{8} - \tfrac{9}{40}\alpha_1 ,$$

hence by (55)

$$\gamma \geqq \tfrac{3}{8} - \tfrac{9}{40}\alpha_1 + \tfrac{9}{10}\alpha_2 \geqq \tfrac{9}{8}\alpha_1 - \tfrac{9}{40}\alpha_1 + \tfrac{9}{10}\alpha_2 = \tfrac{9}{10}(\alpha_1 + \alpha_2).$$

5) $$\tfrac{1}{3} < \alpha_1 < \tfrac{3}{7}.$$

By VI and (52) we have

$$\gamma \geqq \tfrac{1}{3} + \alpha_2 = \tfrac{1}{3} + \tfrac{1}{10}\alpha_2 + \tfrac{9}{10}\alpha_2 \geqq \tfrac{1}{3} + \tfrac{11}{90}\alpha_1 + \tfrac{9}{10}\alpha_2$$
$$> \tfrac{7}{9}\alpha_1 + \tfrac{11}{90}\alpha_1 + \tfrac{9}{10}\alpha_2 = \tfrac{9}{10}(\alpha_1 + \alpha_2).$$

6) $$\alpha_1 \geqq \tfrac{3}{7}.$$

It follows from (52) that $\alpha_2 > \frac{11}{21}$. Because of IV we have

$$\gamma \geqq \frac{\alpha_2}{1 - \alpha_1} > \frac{11 \cdot 7}{21 \cdot 4} > \tfrac{9}{10} > \tfrac{9}{10}(\alpha_1 + \alpha_2).$$

This proves Theorem 6.

THEOREM 7: $\gamma \geqq \alpha_1 + \frac{5}{6}\alpha_2$.

PROOF: It follows from Theorem 6 that

$$\gamma \geqq \tfrac{9}{10}\alpha_1 + \tfrac{9}{10}\alpha_2 \geqq \alpha_1 + \tfrac{5}{6}\alpha_2$$

if $\alpha_2 \geqq \frac{3}{2}\alpha_1$. Therefore we may assume

(56) $$\alpha_2 < \tfrac{3}{2}\alpha_1.$$

2) It follows from II that

$$\gamma \geqq 2\alpha_1 \geqq \alpha_1 + \tfrac{5}{6}\alpha_2$$

if $\alpha_2 \leqq \frac{6}{5}\alpha_1$; hence we may assume

(57) $$\alpha_2 > \tfrac{6}{5}\alpha_1.$$

3) Theorem VI implies

$$\gamma \geqq \alpha_2 + \frac{\alpha_1}{1+\alpha_1} \geqq \alpha_1 + \tfrac{5}{6}\alpha_2$$

if $\alpha_2 \geqq 6\alpha_1^2/(1+\alpha_1)$. Therefore we may assume

$$6\alpha_1^2/(1+\alpha_1) > \alpha_2 > \tfrac{6}{5}\alpha_1$$

because of (57), hence $\alpha_1 > \frac{1}{4}$.

4) Since $\alpha_1 > \frac{1}{4}$, we have

$$\gamma \geqq \alpha_2 + \tfrac{1}{4} \geqq \alpha_1 + \tfrac{5}{6}\alpha_2$$

if $\alpha_2 \geqq 6\alpha_1 - \frac{3}{2}$. Therefore we may assume

$$6\alpha_1 - \tfrac{3}{2} > \alpha_2 > \tfrac{6}{5}\alpha_1$$

because of (57), hence $\alpha_1 > \frac{5}{16}$, $\alpha_2 > \frac{3}{8}$.

5) Now we have $2\alpha_1 + \alpha_2 > 1$, hence by Theorem 5

$$\gamma \geqq \tfrac{1}{2} + \tfrac{1}{2}\alpha_2 \geqq \alpha_1 + \tfrac{5}{6}\alpha_2$$

if $\alpha_2 \leqq \frac{3}{2} - 3\alpha_1$. Therefore we may assume

$$\tfrac{3}{2} - 3\alpha_1 < \alpha_2 < \tfrac{3}{2}\alpha_1$$

because of (56), hence $\alpha_1 > \frac{1}{3}$.

6) Since $\alpha_1 > \frac{1}{3}$, we have

$$\gamma \geqq \alpha_2 + \tfrac{1}{3} \geqq \alpha_1 + \tfrac{5}{6}\alpha_2$$

if $\alpha_1 \leqq \frac{1}{6}\alpha_2 + \frac{1}{3}$. Therefore we may assume

$$\tfrac{1}{6}\alpha_2 + \tfrac{1}{3} < \alpha_1 < \tfrac{5}{6}\alpha_2$$

because of (57), hence $\alpha_2 > \frac{1}{2}$.

7) Since $\alpha_2 > \frac{1}{2}$, it follows for $\alpha_2 \leqq \frac{3}{5}$ that

$$\gamma \geqq \alpha_1 + \tfrac{1}{2} \geqq \alpha_1 + \tfrac{5}{6}\alpha_2 .$$

Hence we may assume $\alpha_2 > \frac{3}{5}$ and $\alpha_1 > \frac{3}{5}$ because of (56). This gives a contradiction to I, and our theorem is proved.

Before improving XI for $n > 2$ we have to prove the following

LEMMA 3: *Suppose* $n = 3$ *and*

$$\alpha_3 \geqq \alpha_2 \geqq \alpha_1 > .24 \tag{58}$$

If γ^*_{123}, d, *and* δ_ν *have the same significance as in Lemma* 2, *then we have* $d \leqq \frac{22}{225}$.

PROOF: Let $\gamma^*_{\kappa\lambda}$ be the lower bound for the density of $A_\kappa + A_\lambda$ ($\kappa \neq \lambda$) obtained by the theorems of this paper. We set

$$d_{\kappa\lambda} = \alpha_\kappa + \alpha_\lambda - \gamma^*_{\kappa\lambda} .$$

Since $\alpha_1 + 2\alpha_2 < 1$, we obtain from (58)

(59) $$\alpha_2 < .38.$$

We have again to distinguish between several cases:

1) $$\tfrac{1}{3} < \alpha_2 < .38.$$

Suppose first $\alpha_1 > \frac{1}{4}$. Then we have by Theorem 1 for $m = 2$

$$d \leqq \delta_1 + \delta_2 \leqq \alpha_1 - \tfrac{1}{4} + \alpha_2 - \tfrac{1}{3} = \alpha_1 + \alpha_2 - \tfrac{7}{12} \leqq \tfrac{2}{3} - \tfrac{7}{12} < \tfrac{22}{225}.$$

Suppose next $\alpha_1 \leqq \frac{1}{4}$; then by (59)

$$d \leqq \delta_1 + \delta_2 < \tfrac{1}{20} + \tfrac{7}{150} = \tfrac{29}{300} < \tfrac{22}{225}.$$

Because of (58) and (59) we may assume

(60) $$.24 < \alpha_2 \leqq \tfrac{1}{3}.$$

2) $$\alpha_3 \geqq \tfrac{857}{1800}.$$

Because of (58) we have $\gamma_{12}^* \geqq 2\alpha_1 > .48$ and

$$\gamma_{123}^* \geqq 2 \min (\gamma_{12}^*, \alpha_3) \geqq 2 \min (\tfrac{48}{100}, \tfrac{857}{1800}) = \tfrac{857}{900}.$$

On the other hand we have by Theorem 3

$$\alpha_1 + \alpha_2 + \alpha_3 < 1.05,$$

hence

$$d < 1.05 - \tfrac{857}{900} = \tfrac{22}{225}.$$

Therefore we may assume

(61) $$\alpha_3 < \tfrac{857}{1800}.$$

3) $$\alpha_2 \leqq \tfrac{1}{4}.$$

a) $$\alpha_3 \leqq \tfrac{1}{3} + \tfrac{79}{900}.$$

Then we have, because of (58),

$$\delta_3 \leqq \max (\tfrac{1}{12}, \tfrac{79}{900}) = \tfrac{79}{900}; \quad d_{12} \leqq \alpha_2 - \alpha_1 \leqq \tfrac{1}{100}; \quad d \leqq d_{12} + \delta_3 \leqq \tfrac{22}{225}.$$

b) $$\alpha_3 > \tfrac{1}{3} + \tfrac{79}{900}.$$

Because of (58) and (61) we have $\gamma_{12}^* \geqq 2\alpha_1 > \alpha_3$, hence $\gamma_{123}^* \geqq 2\alpha_3$, and

$$d \leqq \alpha_1 + \alpha_2 + \alpha_3 - 2\alpha_3 \leqq 2\alpha_2 - \alpha_3 < \tfrac{1}{2} - \tfrac{379}{900} = \tfrac{71}{900} < \tfrac{22}{225}.$$

Therefore the case 3) is also proved and we may assume that $\alpha_2 > \frac{1}{4}$.

4) $$\tfrac{1}{4} < \alpha_2 \leqq \tfrac{1}{4} + \tfrac{43}{900} = \tfrac{67}{225}.$$

We have $\delta_2 \leqq \frac{43}{900}$; $\delta_1 \leqq \frac{1}{20}$; $d \leqq \frac{22}{225}$. Hence we may assume $\alpha_2 > \frac{67}{225}$.

5) $$\tfrac{67}{225} < \alpha_2 \leqq .3.$$

a) $$\alpha_3 \leqq \tfrac{343}{900}.$$

If $\alpha_3 \leqq \frac{1}{3}$, we have $d_{23} \leqq \alpha_3 - \alpha_2 < \frac{32}{900}$. If $\alpha_3 > \frac{1}{3}$, we have $d_{23} \leqq \delta_3 = \frac{343}{900} - \frac{1}{3} = \frac{43}{900}$. Hence

$$d \leqq \delta_1 + d_{23} \leqq \frac{1}{20} + \frac{43}{900} = \frac{22}{225},$$

since $\alpha_1 \leqq \alpha_2 \leqq \frac{3}{10}$. Therefore we assume

(62) $$\alpha_3 > \frac{343}{900}.$$

b) $$\alpha_1 \leqq \frac{1}{4} \quad \text{and} \quad \alpha_3 \leqq \frac{1}{3} + \frac{7}{75}.$$

From (58) and (62) it follows that $\alpha_1 + 2\alpha_3 > 1$. Hence we obtain by Theorem 5

$$\gamma^*_{13} \geqq \frac{1}{2} + \frac{1}{2}\alpha_1$$

and by IV and (58)

$$\gamma^*_{123} \geqq \frac{\gamma^*_{13}}{1-\alpha_2} \geqq \frac{1+\alpha_1}{2(1-\alpha_2)} \geqq \frac{1.24 \cdot 225}{2 \cdot 158} = \frac{279}{316},$$

$$d \leqq \alpha_1 + \alpha_2 + \alpha_3 - \frac{279}{316} < \frac{1}{4} + \frac{3}{10} + \frac{1}{3} + \frac{7}{75} - \frac{279}{316} = \frac{293}{300} - \frac{279}{316} < \frac{22}{225}.$$

c) $$\alpha_1 \leqq \frac{1}{4} \quad \text{and} \quad \alpha_3 > \frac{1}{3} + \frac{7}{75}.$$

We have

$$2\alpha_2 + \alpha_3 > \frac{134}{225} + \frac{1}{3} + \frac{7}{75} > 1,$$

hence by Theorem 5 and by (58)

$$\gamma^*_{123} \geqq \frac{1}{2} + \frac{1}{2}\alpha_3 + \frac{1}{5},$$

$$d = \alpha_1 + \alpha_2 + \alpha_3 - \gamma^*_{123} \leqq \frac{1}{4} + \frac{3}{10} + \frac{857}{3600} - \frac{1}{2} - \frac{1}{5} = \frac{317}{3600} < \frac{22}{225}$$

because of (61).

d) $$\frac{1}{4} < \alpha_1 \leqq \frac{1}{4} + \frac{43}{900}.$$

We have

$$\delta_1 \leqq \frac{43}{900}, \qquad \delta_2 \leqq \frac{1}{20}, \qquad d \leqq \frac{22}{225}.$$

e) $$\alpha_1 > \frac{1}{4} + \frac{43}{900} = \frac{67}{225} \quad \text{and} \quad \alpha_3 \leqq \frac{128}{395} + \frac{22}{225}.$$

It follows similarly as above in 5*b* that

$$\gamma^*_{123} \geqq \frac{1+\alpha_1}{2(1-\alpha_2)} \geqq \frac{73}{79},$$

$$d \leqq \alpha_1 + \alpha_2 + \alpha_3 - \frac{73}{79} < \frac{3}{10} + \frac{3}{10} + \frac{128}{395} + \frac{22}{225} - \frac{73}{79} = \frac{22}{225}.$$

f) $$\alpha_1 > \frac{67}{225} \quad \text{and} \quad \alpha_3 > \frac{128}{395} + \frac{22}{225}.$$

We have

$$2\alpha_2 + \alpha_3 > \frac{134}{225} + \frac{128}{395} + \frac{22}{225} = \frac{52}{75} + \frac{128}{395} > \frac{4}{5}\left(\frac{13}{15} + \frac{2}{5}\right) > 1,$$

hence by Theorem 5

$$\gamma^*_{123} \geqq \tfrac{1}{2} + \tfrac{1}{2}\alpha_3 + \tfrac{1}{4}$$

and by (61)

$$d < \tfrac{3}{5} + \tfrac{857}{3600} - \tfrac{3}{4} = \tfrac{317}{3600} < \tfrac{22}{225}.$$

Therefore the lemma is proved also in the case 5) and we assume

(63) $$\tfrac{3}{10} < \alpha_2 \leqq \tfrac{1}{3}.$$

6) $$\alpha_1 + \alpha_2 \leqq \tfrac{1}{2} + \tfrac{22}{225} = \tfrac{269}{450}.$$

It follows from (63) and (58) that

$$\gamma^*_{12} \geqq \alpha_2 + \tfrac{1}{5} > \tfrac{1}{2};$$

hence, because of VI,

$$\gamma^*_{123} \geqq \alpha_3 + \tfrac{1}{2},$$

$$d = \alpha_1 + \alpha_2 + \alpha_3 - \gamma^*_{123} \leqq \tfrac{1}{2} + \tfrac{22}{225} + \alpha_3 - \alpha_3 - \tfrac{1}{2} = \tfrac{22}{225}.$$

Therefore we may assume

(64) $$\alpha_1 + \alpha_2 > \tfrac{269}{450}$$

and, because of (63),

(65) $$\alpha_1 > \tfrac{269}{450} - \tfrac{1}{3} = \tfrac{119}{450} > \tfrac{1}{4}.$$

7) $$\alpha_3 \leqq \tfrac{1}{3}.$$

a) $$\alpha_1 \geqq \tfrac{64}{225}.$$

By II we have

$$d \leqq \alpha_1 + \alpha_2 + \alpha_3 - 3\alpha_1 \leqq 2\alpha_3 - 2\alpha_1 \leqq \tfrac{2}{3} - \tfrac{128}{225} = \tfrac{22}{225}.$$

b) $$\alpha_1 < \tfrac{64}{225}.$$

Because of (65) we have

$$\delta_1 < \tfrac{31}{900}, \qquad d_{23} \leqq \alpha_3 - \alpha_2 \leqq \tfrac{1}{3} - \tfrac{3}{10} = \tfrac{1}{30}, \qquad d < \tfrac{61}{900} < \tfrac{22}{225}.$$

Therefore we may assume $\alpha_3 > \frac{1}{3}$.

8) $$\tfrac{1}{3} < \alpha_3 \leqq \tfrac{1}{3} + \tfrac{101}{1800} = \tfrac{701}{1800}.$$

For $\alpha_1 \leqq \frac{1}{4} + \frac{1}{24}$ we have by (65)

(66) $$d_{12} \leqq \delta_1 \leqq \tfrac{1}{24}.$$

For $\alpha_1 > \frac{1}{4} + \frac{1}{24}$ the inequality (66) is also valid, since we have

$$d_{12} \leqq \alpha_2 - \alpha_1 \leqq \tfrac{1}{3} - \tfrac{7}{24} = \tfrac{1}{24}.$$

Hence we obtain

$$d \leqq d_{12} + \delta_3 \leqq \tfrac{1}{24} + \tfrac{101}{1800} = \tfrac{22}{225}$$

and we may assume

$$\alpha_3 > \tfrac{701}{1800}. \tag{67}$$

9) It follows from (66) and IV that

$$\gamma^*_{12} \geqq \alpha_1 + \alpha_2 - \tfrac{1}{24},$$

$$\gamma^*_{123} \geqq (\alpha_1 + \alpha_2 - \tfrac{1}{24})/(1 - \alpha_3),$$

$$\begin{aligned} d &\leqq \alpha_1 + \alpha_2 + \alpha_3 - \frac{\alpha_1 + \alpha_2 - \frac{1}{24}}{1 - \alpha_3} \\ &= \alpha_3 - \frac{\alpha_3(\alpha_1 + \alpha_2) - \frac{1}{24}}{1 - \alpha_3} \leqq \alpha_3 - (\tfrac{269}{450}\alpha_3 - \tfrac{1}{24})(1 - \alpha_3)^{-1} \end{aligned} \tag{68}$$

because of (64). If we write

$$f(\alpha_3) = \alpha_3 - (\tfrac{269}{450}\,\alpha_3 - \tfrac{1}{24})(1 - \alpha_3)^{-1},$$

we have

$$f'(\alpha_3) = 1 - \{\tfrac{269}{450}\,(1 - \alpha_3) + \tfrac{269}{450}\,\alpha_3 - \tfrac{1}{24}\}(1 - \alpha_3)^{-2} = 1 - \tfrac{1001}{1800}\,(1 - \alpha_3)^{-2}.$$

Therefore $f(\alpha_3)$ has its maximum in the interval $0 \leqq \alpha_3 < 1$ for

$$\alpha_3 = 1 - (\tfrac{1001}{1800})^{\frac{1}{2}} < 1 - \tfrac{31}{43} < \tfrac{701}{1800}.$$

Hence $f(\alpha_3)$ is decreasing in the interval $\frac{701}{1800} < \alpha_3 < 1$ and it follows from (67) and (68) that

$$\begin{aligned} d &\leqq \frac{701}{1800} - \left(\frac{701}{1800} \cdot \frac{269}{450} - \frac{1}{24}\right) \cdot \frac{1800}{1099} \\ &= \frac{701 \cdot 1099 - 701 \cdot 1076 + 75 \cdot 1800}{1800 \cdot 1099} < \frac{701 + 5400}{1800 \cdot 40} < \frac{22}{225}. \end{aligned}$$

This proves Lemma 3.

Using this lemma we can improve XI.

THEOREM 8: *Write* $\alpha_1 + \alpha_2 + \cdots + \alpha_n = \sigma_n$; *then we have*

$$\gamma \geqq \tfrac{9}{10}\,\sigma_2; \qquad \gamma \geqq \tfrac{81}{92}\,\sigma_3 > .8804\sigma_3; \qquad \gamma \geqq \tfrac{972}{1117}\,\sigma_4 > .8701\sigma_4;$$

$$\gamma \geqq \tfrac{340\,20}{391\,97}\,\sigma_n > .8679\sigma_n \qquad \textit{for } n \geqq 5.$$

PROOF: The case $n = 2$ is already proved by Theorem 6.

1) $n = 3$

a) $\alpha_1 \leqq \tfrac{6}{25}.$

Denote by α_{23} the density of $A_2 + A_3$. It follows from Theorem 6 that

$$\alpha_{23} \geqq \tfrac{9}{10}(\alpha_2 + \alpha_3),$$

hence from Theorem VII for $k = 6$ and from II

$$\gamma \geqq \tfrac{9}{10}(\alpha_2 + \alpha_3) + \tfrac{5}{6}\alpha_1, \tag{69}$$

$$\gamma \geqq 3\alpha_1, \tag{70}$$

and, because of (69) and (70),

$$90\gamma + 2\gamma \geqq 81(\alpha_2 + \alpha_3) + 75\alpha_1 + 6\alpha_1,$$

$$\gamma \geqq \tfrac{81}{92}\sigma_3.$$

b) $\alpha_1 > \tfrac{6}{25}$.

Suppose

$$\gamma < \tfrac{81}{92}\sigma_3. \tag{71}$$

Then we have

$$\tfrac{18}{25} < 3\alpha_1 \leqq \gamma < \tfrac{81}{92}\sigma_3,$$

$$\sigma_3 > \tfrac{184}{225}. \tag{72}$$

On the other hand it follows from Lemma 3 and (71) that

$$\sigma_3 - \tfrac{22}{225} \leqq \gamma^*_{123} \leqq \gamma < \tfrac{81}{92}\sigma_3,$$

$$\tfrac{11}{92}\sigma_3 < \tfrac{22}{225},$$

$$\sigma_3 < \tfrac{184}{225}.$$

This is incompatible with (72). Therefore the supposition (71) is impossible and in both cases we have

$$\gamma \geqq \tfrac{81}{92}\sigma_3. \tag{73}$$

2) $n = 4$.

It is possible to treat this case as we have treated the case $n = 3$. But the proof of the lemma analogous to the Lemma 3 is troublesome. Therefore we use a method which is simpler, but less sharp. Denote by α_{234} the density of $A_2 + A_3 + A_4$. It follows from (73) that

$$\alpha_{234} \geqq \tfrac{81}{92}(\alpha_2 + \alpha_3 + \alpha_4),$$

hence from VIII and from II

$$\gamma \geqq \tfrac{81}{92}(\alpha_2 + \alpha_3 + \alpha_4) + \tfrac{5}{6}\alpha_1, \tag{74}$$

$$\gamma \geqq 4\alpha_1. \tag{75}$$

Therefore we have by (74) and (75)

$$1104\gamma + 13\gamma \geqq 972\,(\alpha_2 + \alpha_3 + \alpha_4) + 920\alpha_1 + 52\alpha_1,$$

$$\gamma \geqq \tfrac{972}{1117}\,\sigma_4.$$

3) $$n = 5.$$

In similar way we obtain

$$39095\gamma + 102\gamma \geqq 39095\,\{\tfrac{972}{1117}\,(\alpha_2 + \alpha_3 + \alpha_4 + \alpha_5) + \tfrac{6}{7}\,\alpha_1\} + 510\alpha_1,$$

$$\gamma \geqq \tfrac{340\,20}{391\,97}\,\sigma_5.$$

4) $$n > 5.$$

It follows from VIII that

$$\gamma \geqq \tfrac{340\,20}{391\,97}\,(\alpha_n + \alpha_{n-1} + \alpha_{n-2} + \alpha_{n-3} + \alpha_{n-4}) + \tfrac{7}{8}\,\alpha_{n-5} + \cdots + \frac{n+1}{n+2}\,\alpha_1,$$

hence, because of $\tfrac{340\,20}{391\,97} < \tfrac{7}{8}$,

$$\gamma \geqq \tfrac{340\,20}{391\,97}\,\sigma_n.$$

7. Estimates for the sum of the densities for $n \geqq 5$

In order to obtain an estimate for the sum of the densities for $n = 5$ we could apply a similar method as in 5). But the proof of the lemma analogous to the Lemmas 1 and 2 would be very complicate. Therefore we are content to prove the following

THEOREM 9: *Suppose that $A_1 + A_2 + \cdots + A_n$ is not the set of all positive integers. Then we have*

$$\sigma_5 < \tfrac{43}{39} < 1.1026 \text{ and } \sigma_n < \frac{39197}{34020n + 5177} < 1.1522 \quad \text{for } n > 5.$$

PROOF: For $n = 5$ we obtain from Theorem 2 for $k = 2$ and $l = 1$

(76) $$\alpha_{12} \geqq \alpha_2 + \tfrac{6}{7}\,\alpha_1$$

and from II and I

(77) $$\alpha_{345} \geqq 3\alpha_3,$$

(78) $$\alpha_{12} + \alpha_{345} < 1,$$

hence, because of (76), (77), and (78)

(79) $$\tfrac{6}{7}\,\alpha_1 + \alpha_2 + 3\alpha_3 < 1.$$

Moreover we have

(80) $$\alpha_1 + 4\alpha_2 < 1,$$

(81) $$2\alpha_1 + 3\alpha_3 < 1,$$

and finally by Theorem 2

$$\alpha_{1234} \geqq \alpha_4 + \tfrac{4}{5}\alpha_3 + \tfrac{5}{6}\alpha_2 + \tfrac{6}{7}\alpha_1,$$

hence, because of I

$$\alpha_5 + \alpha_4 + \tfrac{4}{5}\alpha_3 + \tfrac{5}{6}\alpha_2 + \tfrac{6}{7}\alpha_1 \leqq \alpha_5 + \alpha_{1234} < 1. \tag{82}$$

If we multiply (79), (80), (81), (82) by $\frac{3}{130}$, $\frac{14}{390}$, $\frac{17}{390}$, 1 respectively and add these inequalities, we obtain

$$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 < \tfrac{3}{130} + \tfrac{14}{390} + \tfrac{17}{390} + 1 = \tfrac{43}{39} < 1.1026.$$

For $n > 5$ we are content to prove an inequality less sharp. It follows from Theorem 8 and from II that

$$\tfrac{34020}{39197}(\alpha_1 + \alpha_2 + \cdots + \alpha_{n-1}) + \alpha_n < 1,$$

and because of $\alpha_n \geqq (\alpha_1 + \alpha_2 + \cdots + \alpha_n)/n$

$$\frac{34020}{39197}(\alpha_1 + \alpha_2 + \cdots + \alpha_{n-1} + \alpha_n) + \frac{5177}{39197}\,\frac{\alpha_1 + \alpha_2 + \cdots + \alpha_n}{n} < 1,$$

$$\alpha_1 + \alpha_2 + \cdots + \alpha_n < \frac{39197n}{34020n + 5177} < 1.1522.$$

8. Improvement of a theorem of Schur

THEOREM 10: *If we set* $c_n^{(1)} = n$ *and for* $m > 1$

$$c_n^{(m)} = \frac{n(n+2)c_{n-1}^{(m-1)}}{(n+2)c_{n-1}^{(m-1)} + n^2 + n - 1} \qquad \text{if } c_{n-1}^{(m-1)} > \frac{n+1}{n+2}, \tag{83}$$

$$c_n^{(m)} = c_{n-1}^{(m-1)} \qquad \text{if } c_{n-1}^{(m-1)} \leqq \frac{n+1}{n+2}, \tag{84}$$

then we have

$$\gamma \geqq c_n^{(m)}(\alpha_1 + \alpha_2 + \cdots + \alpha_m).$$

PROOF: For $m = 1$ the assertion follows from II. Therefore we may assume that the theorem is proved for $m - 1$ and every n. Denoting the density of $A_2 + A_3 + \cdots + A_n$ by β we have

$$\beta \geqq c_{n-1}^{(m-1)}(\alpha_2 + \alpha_3 + \cdots + \alpha_m). \tag{85}$$

We have now to distinguish between two cases:

a) $$c_{n-1}^{(m-1)} \leqq \frac{n+1}{n+2}.$$

It follows from VIII and from (85) that

$$\gamma \geqq \beta + \frac{n+1}{n+2}\alpha_1 \geqq c_{n-1}^{(m-1)}(\alpha_2 + \alpha_3 + \cdots + \alpha_m) + \frac{n+1}{n+2}\alpha_1$$

$$\geqq c_{n-1}^{(m-1)}(\alpha_1 + \alpha_2 + \cdots + \alpha_m),$$

hence, because of (84),

$$\gamma \geqq c_n^{(m)}(\alpha_1 + \alpha_2 + \cdots + \alpha_m).$$

b) $$c_{n-1}^{(m-1)} > \frac{n+1}{n+2}.$$

If we set

$$a = \frac{n(n+2)}{c_{n-1}^{(m-1)}(n+2) + n^2 + n - 1},$$

we have

$$a < \frac{n(n+2)}{n+1+n^2+n-1} = 1.$$

Therefore applying IX we obtain

$$\gamma \geqq a\beta + \alpha_1\left\{n - \frac{n^2+n-1}{n+2}\,a\right\}$$

$$= \frac{n(n+2)\beta}{c_{n-1}^{(m-1)}(n+2) + n^2 + n - 1} + \alpha_1\left\{n - \frac{n(n^2+n-1)}{c_{n-1}^{(m-1)}(n+2) + n^2 + n - 1}\right\}$$

$$= \frac{n}{c_{n-1}^{(m-1)}(n+2) + n^2 + n - 1}\{(n+2)\beta + \alpha_1 c_{n-1}^{(m-1)}(n+2)\}$$

$$\geqq \frac{n(n+2)c_{n-1}^{(m-1)}}{c_{n-1}^{(m-1)}(n+2) + n^2 + n - 1}(\alpha_1 + \alpha_2 + \cdots + \alpha_m)$$

because of (85). Hence Theorem 10 is proved.

COROLLARY: *If $A_1 + A_2 + \cdots + A_n$ is not the sum of all positive integers, we have*

$$\alpha_1 + \alpha_2 + \cdots + \alpha_m < 1/c_n^{(m)}.$$

The constants $c_n^{(m)}$ can be determined successively by (83) and (84). We have $c_n^{(1)} = n$,

$$c_n^{(2)} = \frac{n}{2 + 1/(n-1)(n+2)},$$

and so on.

We want to obtain estimates for $c_n^{(m)}$. To this end we prove

THEOREM 11: *Set $n - m = d$ for $n > m$. For a fixed value of d we denote by z the greatest integer such that*

$$\log(z-1) + \frac{1}{z+1} \leqq d + \log d + 2. \tag{86}$$

Then we have

$$c_n^{(n-d)} > \frac{n+1}{n-d+\log n - \log d} \qquad \text{for } d+2 \leqq n \leqq z, \tag{87}$$

$$c_n^{(n-d)} > \frac{z+1}{z+2} > \frac{e^{d+1}d+1}{e^{d+1}d+2} \qquad \text{for } n \geqq z. \tag{88}$$

PROOF: First we assume that

$$c_{n-1}^{(m-1)} > \frac{n+1}{n+2}. \tag{89}$$

We maintain that then

$$c_{n-2}^{(m-2)} > \frac{n}{n+1}. \tag{90}$$

For, if $c_{n-2}^{(m-2)} \leqq n/(n+1)$ it follows from (84) that

$$c_{n-1}^{(m-1)} = c_{n-2}^{(m-2)} \leqq \frac{n}{n+1} < \frac{n+1}{n+2}.$$

This contradicts (89). Therefore (90) is proved; we infer by induction

$$c_{n-\mu}^{(m-\mu)} > \frac{n-\mu+2}{n-\mu+3} \qquad \text{for } \mu = 2, 3, \cdots, m-1 \tag{91}$$

if (89) is satisfied. Because of (91) we obtain $c_{n-m+2}^{(2)}, c_{n-m+3}^{(3)}, \cdots, c_n^{(m)}$ successively from (83).

We set $y_n^{(m)} = 1/c_n^{(m)}$. It follows from (83) that

$$n y_n^{(m)} = \frac{n^2+n-1}{n+2} y_{n-1}^{(m-1)} + 1 < \frac{n^2}{n+1} y_{n-1}^{(m-1)} + 1,$$

hence

$$(n+1) y_n^{(m)} < n y_{n-1}^{(m-1)} + \frac{n+1}{n}.$$

Applying this formula successively we obtain

$$\begin{aligned}
(n+1)y_n^{(m)} &< (n-1)y_{n-2}^{(m-2)} + \frac{n}{n-1} + \frac{n+1}{n} < \cdots \\
&< (n-m+2)y_{n-m+1}^{(1)} + \frac{n-m+3}{n-m+2} + \frac{n-m+4}{n-m+3} + \cdots + \frac{n+1}{n} \\
&= \frac{n-m+2}{n-m+1} + \frac{n-m+3}{n-m+2} + \cdots + \frac{n+1}{n} \\
&= m + \sum_{\nu=n-m+1}^{n} \frac{1}{\nu} < m + \int_{n-m}^{n} \frac{dx}{x} = m + \log \frac{n}{n-m}.
\end{aligned}$$

Therefore we have

$$c_n^{(m)} > \frac{n+1}{m+\log n-\log(n-m)} \tag{92}$$

if (89) is satisfied.

Next we want to prove that (89) is satisfied if

$$d+2 \leqq n \leqq z$$

where z is defined by (86). To this end we prove that

$$c_{d+r}^{(r)} > \frac{d+r+1}{r+\log(d+r)-\log d} \geqq \frac{d+r+2}{d+r+3} \tag{93}$$

if the condition

$$d+2 \leqq d+r \leqq z-1 \tag{94}$$

is satisfied.

We have

$$c_{d+1}^{(1)} = d+1 > (d+3)/(d+4),$$

hence it follows from (92) that

$$c_{d+2}^{(2)} > \frac{d+3}{2+\log(d+2)-\log d} > 1 > \frac{d+4}{d+5}$$

since

$$d+3 \geqq 4 > 2+\log 3 \geqq 2+\log\left(1+\frac{2}{d}\right) = 2+\log(d+2)-\log d.$$

Therefore (93) is proved for $r=2$. We assume it to be proved by induction for $r = 2, 3, \cdots, \rho-1$ where

$$d+\rho-1 \leqq z-2. \tag{95}$$

We then have

$$c_{d+\rho-1}^{(\rho-1)} > \frac{d+\rho}{\rho-1+\log(d+\rho-1)-\log d} \geqq \frac{d+\rho+1}{d+\rho+2}.$$

Hence the condition (89) is satisfied for $n = d+\rho$ and it follows from (92) that

$$c_{d+\rho}^{(\rho)} > \frac{d+\rho+1}{\rho+\log(d+\rho)-\log d}. \tag{96}$$

On the other hand, it follows from (95) that $d+\rho+1 \leqq z$ and therefore from (86) that

$$\log(d+\rho)+\frac{1}{d+\rho+2} \leqq d+\log d+2.$$

Consequently we have

$$d+\rho+2-\frac{1}{d+\rho+2} \geqq \rho+\log(d+\rho)-\log d,$$

$$\frac{(d+\rho+1)(d+\rho+3)}{d+\rho+2} \geqq \rho+\log(d+\rho)-\log d,$$

hence by (96)

$$c_{d+\rho}^{(\rho)} > \frac{d+\rho+1}{\rho+\log(d+\rho)-\log d} \geqq \frac{d+\rho+2}{d+\rho+3}.$$

Therefore (93) is proved also for $r=\rho$ and hence by induction for each r satisfying (94). It follows from (93) for $d+r=z-1$ that (89) is true for $n=z$. This proves (87) because of (92).

Now we have to prove (88). Let s be the greatest integer such that

$$c_{z-1+s}^{(z-1+s-d)} > \frac{z+s+1}{z+s+2};$$

then we have $s \geqq 0$ because of (93). It follows similarly as above that

$$c_{z-1+\sigma}^{(z-1+\sigma-d)} > \frac{z+\sigma+1}{z+\sigma+2} \geqq \frac{z+1}{z+2} \qquad \text{for } \sigma = 0, 1, \cdots, s. \tag{97}$$

Furthermore we have

$$c_{z-1+\sigma}^{(z-1+\sigma-d)} \leqq \frac{z+\sigma+1}{z+\sigma+2} \qquad \text{for } \sigma \geqq s+1,$$

hence by (84)

$$c_{z-1+\sigma}^{(z-1+\sigma-d)} = c_{z-1+\sigma-1}^{(z-1+\sigma-1-d)} = \cdots = c_{z+s}^{z+s-d} \qquad \text{for } \sigma \geqq s+2. \tag{98}$$

Finally we have to consider the case $n=z+s$. Denoting the density of $A_2+A_3+\cdots+A_{s+s}$ by β we obtain from VIII

$$\gamma \geqq \beta + \frac{z+s+1}{z+s+2}\alpha_1 \geqq c_{z+s-1}^{(z+s-1-d)}(\alpha_2+\alpha_3+\cdots+\alpha_{z+s-d}) + \frac{z+s+1}{z+s+2}\alpha_1$$

$$> \frac{z+1}{z+2}(\alpha_1+\alpha_2+\cdots+\alpha_{z+s-d})$$

because of (97); therefore we have

$$c_{z+s}^{(z+s-d)} > \frac{z+1}{z+2}. \tag{99}$$

It follows from (98) and (99) that

$$c_{z-1+\sigma}^{(z-1+\sigma-d)} > \frac{z+1}{z+2} \qquad \text{for } \sigma \geqq s+1. \tag{100}$$

Moreover, the condition (86) implies

$$\log z + \frac{1}{z+2} > d + \log d + 2,$$
$$z > e^{d+1}d. \tag{101}$$

Now (88) follows from (97) and (100). Hence Theorem 11 is proved.

The inequality (88) can be improved a little. Since $d \geqq 1$, it follows from (101) that

$$\log z + \frac{1}{z+2} > 3,$$
$$z \geqq 20, \tag{102}$$

hence we obtain from (101)

$$\log z > d + \log d + \tfrac{43}{22},$$
$$c_n^{(n-d)} > \frac{z+1}{z+2} > \frac{e^{d+43/22}d+1}{e^{d+43/22}d+2} \qquad \text{if } n \geqq z.$$

I. Schur has proved that

$$\gamma \geqq g_n^{(m)}(\alpha_1 + \alpha_2 + \cdots + \alpha_m)$$

where the constants $g_n^{(m)}$ are defined by

$$g_n^{(1)} = n, \qquad g_n^{(m)} = \frac{n}{1 + (n^2-n)^{\frac{1}{2}}/g_{n-1}^{(m-1)}} \qquad \text{for } m > 1. \tag{103}$$

We want to prove that $c_n^{(m)} \geqq g_n^{(m)}$. For $m = 1$ we have $c_n^{(1)} = g_n^{(1)} = n$. Furthermore we have for $n \geqq 2$

$$(n^2-n)(n+2)^2 = n^4 + 3n^3 - 4n > n^4 + 2n^3 - n^2 - 2n + 1 = (n^2+n-1)^2,$$

hence

$$(n^2-n)^{\frac{1}{2}} > \frac{n^2+n-1}{n+2}.$$

Therefore it follows from (103) that we have $c_n^{(m)} \geqq g_n^{(m)}$ for all those $c_n^{(m)}$ which are defined by (83).

From the proof of Theorem 11 we deduce that we have to apply (83) certainly as long as $n \leqq z$ for a fixed value of d. Therefore we may assume now that

$$n \geqq z + 1 > e^{d+1}d + 1.$$

It follows from (88) that it is sufficient to prove that

$$\frac{e^{d+1}d+1}{e^{d+1}d+2} \geqq g_n^{(n-d)} \qquad \text{for } n > e^{d+1}d + 1. \tag{104}$$

From (103) Schur obtained

$$g_n^{(n-d)} = \frac{n^{\frac{1}{2}}}{(d+1)^{-\frac{1}{2}} + (d+2)^{-\frac{1}{2}} + \cdots + n^{-\frac{1}{2}}}. \tag{105}$$

We have

$$\begin{cases} (d+1)^{-\frac{1}{2}} + (d+2)^{-\frac{1}{2}} + \cdots + n^{-\frac{1}{2}} > \int_{d+1}^{n+1} x^{-\frac{1}{2}}\, dx \\ \qquad = 2(n+1)^{\frac{1}{2}} - 2(d+1)^{\frac{1}{2}} > 2n^{\frac{1}{2}} - 2(d+1)^{\frac{1}{2}}. \end{cases} \tag{106}$$

Because of (104), (105), and (106), it is sufficient to prove that

$$\frac{e^{d+1}d + 1}{e^{d+1}d + 2} > \frac{n^{\frac{1}{2}}}{2n^{\frac{1}{2}} - 2(d+1)^{\frac{1}{2}}}.$$

This is satisfied if

$$e^{d+1}dn^{\frac{1}{2}} > 2(e^{d+1}d + 1)(d+1)^{\frac{1}{2}},$$
$$n^{\frac{1}{2}} > 3(d+1)^{\frac{1}{2}},$$
$$n > 9(d+1).$$

This, however, is true for $d = 1$ because of (102), and for $d > 1$ because of (104). Hence we have $c_n^{(m)} \geqq g_n^{(m)}$ and Theorem 10 is sharper than Schur's theorem.

THEOREM 12: *Let ϵ be any positive quantity. For every fixed m and all sufficiently large n we have*

$$\frac{n}{m} - c_n^{(m)} < \epsilon.$$

PROOF: Let m be fixed. We choose n so large that the following conditions are satisfied

$$\log(n-1) + \frac{1}{n+1} < n - m + \log(n-m) + 2 \tag{107}$$

and

$$n > \text{Max}\,(2m, 2/\epsilon). \tag{108}$$

Then it follows from (107), (86), and (87) that

$$c_n^{(m)} > \frac{n+1}{m + \log n - \log(n-m)},$$

$$\frac{n}{m} - c_n^{(m)} < \frac{n}{m} - \frac{n+1}{m + \log(1 + m/(n-m))}.$$

Because of (108) we have $m/(n-m) < 1$, hence

$$\frac{n}{m} - c_n^{(m)} < \frac{n}{m} - \frac{n+1}{m + m/(n-m)}$$

$$= \frac{n}{m} - \frac{(n+1)(n-m)}{m(n-m+1)} = \frac{1}{n-m+1} < \frac{2}{n} < \epsilon.$$

INSTITUTE FOR ADVANCED STUDY
PRINCETON, N. J.

ANNALS OF MATHEMATICS
Vol. 42, No. 4, October, 1941

ON SOME ASYMPTOTIC FORMULAS IN THE THEORY OF THE "FACTORISATIO NUMERORUM"

BY P. ERDÖS

(Received December 2, 1940)

Let $1 < a_1 \leqq a_2 \leqq \cdots$ be a sequence of integers. Denote by $f(n)$ the number of representations of n as the product of the a's, where two representations are considered equal only if they contain the same factors in the same order. As far as I know the first papers written on the subject are those of L. Kalmár,[1] who proved by using the methods of analytic number theory that if $a_k = k + 1$ then

$$F(n) = \sum_{r=1}^{n} f(r) = -\frac{n^\rho}{\rho\zeta'(\rho)}[1 + o(1)], \tag{1}$$

ρ is defined as the unique positive root of $\zeta(\rho) = 2$. He also gives estimates for the error term.

Another paper on this subject is that of E. Hille.[2] He obtains among others the following results: Let $p_1 < p_2 < \cdots$ be a sequence of primes and $a_1 < a_2 < \cdots$ the sequence of integers composed of these primes, then

$$F(n) = cn^\rho[1 + o(1)], \tag{2}$$

where $\sum_i \frac{1}{a_i^\rho} = 1$, $\rho > 0$. Hille uses the theorem of Wiener and Ikehara.

In the present paper we assume that $\sum \frac{1}{a_i^{1+\epsilon}}$ converges for every ϵ and that the a's are not all powers of a_1, then we prove that

$$F(n) = cn^\rho[1 + o(1)], \tag{3}$$

where $\sum_i \frac{1}{a_i^\rho} = 1$, $\rho > 0$. The proof will be elementary.

First we need 2 Lemmas.

LEMMA 1

$$F(n) = \sum_k F\left[\frac{n}{a_k}\right] + 1.^3 \tag{4}$$

PROOF. Follows immediately by considering those products in which a_k is the first factor, and summing for a_k.

[1] L. Kalmár, Acta Litt ac Scient. Szeged, Tom. 5 (1930) p. 95–107.
[2] E. Hille, Acta Arithmetica Vol. 2 (1937) p. 134–146.
[3] The use of this identity was suggested to me by L. Kalmár.

LEMMA 2.

$$0 < \underline{\lim} \frac{F(n)}{n^\rho} \leqq \overline{\lim} \frac{F(n)}{n^\rho} < \infty. \tag{5}$$

PROOF. Put $F(n) = c_n n^\rho$. We have from (4)

$$c_n n^\rho < \max_{i \leqq \frac{n}{2}} c_i \sum_{a_k} \frac{n^\rho}{a_k^\rho} + 1,$$

hence

$$c_n < \max_{i \leqq \frac{n}{2}} c_i + \frac{1}{n^\rho}.$$

Thus by induction

$$c_n < 1 + \sum_{2^{m-1} < n} \frac{1}{2^{m\rho}} < \infty,$$

which proves the first half of (5).

The proof of the second half of (5) will be slightly more complicated. Put $F(n) = c_n'(n+1)^\rho$. It suffices to prove that $\underline{\lim}\, c_n' > 0$. From $\left[\frac{n}{a_k}\right] \geqq \frac{n+1}{a_k} - 1$ we obtain by (4)

$$c_n'(n+1)^\rho > \min_{i \leqq \frac{n}{2}} c_i' \sum_{a_k \leqq n} \frac{(n+1)^\rho}{a_k^\rho} = \min_{i \leqq \frac{n}{2}} c_i'(n+1)^\rho \left(1 - \sum_{a_k > n} \frac{1}{a_k^\rho}\right).$$

Thus

$$c_n' > \min_{i \leqq \frac{n}{2}} c_i' \left(1 - \sum_{a_k > n} \frac{1}{a_k^\rho}\right).$$

Hence by induction

$$c_n' > \prod_{2^{m-1} < n} \left(1 - \sum_{a_k > 2^m} \frac{1}{a_k^\rho}\right).$$

The product on the right side (if extended to infinity) converges since

$$\sum_{m=1}^{\infty} \sum_{a_k > 2^m} \frac{1}{a_k^\rho} \leqq \sum_{a_k} \frac{\log a_k}{a_k^\rho} < c \sum \frac{1}{a_k^{1+\varepsilon}}$$

converges. This proves $\underline{\lim}\, c_n' > 0$, and completes the proof of Lemma 2.

Now we can prove our theorem. Suppose that (3) does not hold, denote

$$0 < c = \underline{\lim} \frac{F(n)}{n^\rho} = \underline{\lim} \frac{F(n)}{(n+1)^\rho} < \overline{\lim} \frac{F(n)}{n^\rho} = \overline{\lim} \frac{F(n)}{(n+1)^\rho} = C < \infty. \tag{6}$$

Let m be sufficiently large and such that $F(m) > (C - \delta)(m + 1)^\rho$. Clearly a fixed k exists (depending only on c and C) such that for every x satisfying $m \leqq x \leqq m(1 + k)$

$$\frac{F(x)}{(x+1)^\rho} > \frac{C+c}{2}. \tag{7}$$

Now let a_i be the least a which is not a power of a_1. Consider any x satisfying $ma_1 \leqq x \leqq ma_1(1 + k)$. We have by (4), (6), (7) and $\left[\frac{x}{a_i}\right] + 1 \geqq \frac{x+1}{a_i}$

$$F(x) > \sum_{a_i \leqq x} F\left[\frac{x}{a_i}\right] > \frac{c+C}{2}\frac{(x+1)^\rho}{a_1^\rho} + c\sum_{a_i > a_1}\frac{(x+1)^\rho}{a_i^\rho} - o(x^\rho). \tag{8}$$

Thus

$$\frac{F(x)}{(x+1)^\rho} > c + \frac{C-c}{2a_1^\rho} - o(1). \tag{9}$$

Similarly we obtain that for the x satisfying $a_1^\alpha a_i^\beta m \leqq x \leqq a_1^\alpha a_i^\beta m(1 + k)$

$$\frac{F(x)}{(x+1)^\rho} > c + \delta_{\alpha,\beta}, \tag{10}$$

where $\delta_{\alpha,\beta}$ depends only upon α and β. It is well known that the quotient of two consecutive integers of the form $a_1^\alpha a_i^\beta$ tends to 1. Thus there exists a sequence of integers $A_1 < A_2 < \cdots < A_r$ all of the form $a_1^\alpha a_i^\beta$ and satisfying

$$\frac{A_{i+1}}{A_i} < 1 + k, \qquad i = 1, 2, \cdots r - 1 \quad \text{and} \quad A_r > a_1 A_1.$$

Thus by (10) and since the intervals $[A_i m, A_i m(1 + k)]$ and $[A_{i+1}m, A_{i+1}m(1 + k)]$ overlap we have for $A_1 m \leqq x \leqq a_1 A_1 m$

$$\frac{F(x)}{(x+1)^\rho} > c + \min \delta_{\alpha,\beta} = c + \delta, \tag{11}$$

for sufficiently large m, where δ is fixed and depends only on c and C. Consider now the integers x satisfying $a_1 A_1 m \leqq x \leqq a_1^2 A_1 m_1$ by (4), (6) and (11) we obtain as in (8) and (9)

$$\frac{F(x)}{(x+1)^\rho} > (c+\delta)\frac{1}{a_1^\rho} + c\sum_{a_i > a_1}\frac{1}{a_i^\rho} - o(1) = c + \delta\left(1 - \sum_{a_i > a_1}\frac{1}{a_i^\rho}\right) - o(1).$$

$\left(\text{i.e. } \frac{x}{a_1} \text{ lies in } [A_1 m, A_1 m(1 + k)]\right)$. Similarly for the integers satisfying $a_1^2 A_1 m \leqq x \leqq a_2^3 A_1 m$ we have

$$\frac{F(x)}{(x+1)^\rho} > \left[c + \delta\left(1 - \sum_{a_i > a_1}\frac{1}{a_i^\rho}\right)\right]\sum_{a_i \leqq a_2^2}\frac{1}{a_i^\rho} + c\sum_{a_i > a_2^2}\frac{1}{a_i^\rho}$$

$$- o(1) > c + \delta\left(1 - \sum_{a_i > a_1}\frac{1}{a_i^\rho}\right)\left(1 - \sum_{a_i > a_1^2}\frac{1}{a_i^\rho}\right) - o(1).$$

Finally we obtain for $a_1^{k-1}A_1m \leqq x \leqq a_1^kA_1m$ (k fixed, m sufficiently large)

$$\frac{F(x)}{(x+1)^\rho} > c + \delta \prod_{r=1}^{k}\left(1 - \sum_{a_i > a_1^r} \frac{1}{a_i^\rho}\right) - o(1). \tag{12}$$

Denote

$$\prod_{r=1}^{\infty}\left(1 - \sum_{a_i > a_1^r} \frac{1}{a_i^\rho}\right) = \eta.$$

The product converges since $\sum \frac{\log a_i}{a_i^\rho}$ converges. From (12) we have for $A_1m \leqq x \leqq a_1^kA_1m$

$$\frac{F(x)}{(x+1)^\rho} > c + \frac{\delta\eta}{2}. \tag{13}$$

Now choose k so great that

$$\prod_{r>k} \sum_{a_i \leqq a_1^r} \frac{1}{a_i^\rho} > \frac{c + \frac{1}{4}\delta\eta}{c + \frac{1}{2}\delta\eta}. \tag{14}$$

Then from (13) and (4) we have for $A_1a_1^km \leqq x \leqq A_1a_1^{k+1}m$

$$F(x) > \sum_{a_i \leqq a_1^{k+1}} F\left[\frac{x}{a_i}\right] > \left(c + \frac{\delta\eta}{2}\right) \sum_{a_i \leqq a_1^{k+1}} \frac{(x+1)^\rho}{a_i^\rho}.$$

Similarly for any r, in the interval $A_1a_1^rm \leqq x \leqq A_1a_1^{r+1}m$ we have by (14)

$$\frac{F(x)}{(x+1)^\rho} > \left(c + \frac{\delta\eta}{2}\right) \prod_{l>k} \sum_{a_i < a_l} \frac{(x+1)}{\rho} > \frac{c + \delta\eta}{4}.$$

Thus $\lim \frac{F(x)}{(x+1)^\rho} > c$. This contradicts (6) and completes the proof of our theorem.

It is easy to see that in our theorem, we can replace the assumption that $\sum \frac{1}{a_i^{1+\epsilon}}$ converges by the following slightly more general one: There exists a $k > 0$ such that $\sum \frac{1}{a_i^k}$ converges, and $\sum \frac{\log a_i}{a_i^k}$ converges too.

Let $a_k = k + 1$. By using Lemma 2 we can prove that constants c_1 and c_2 exist, $0 < c_2 < c_1 < 1$, such that for infinitely many n

$$f(n) > \frac{n^\rho}{e^{(\log n)^{c_1}}}$$

and that for all $n > n_0$

$$f(n) < \frac{n^\rho}{e^{(\log n)^{c_2}}}.$$ [4]

As I shall show in another paper the methods used here yield some asymptotic formulas in the theory of partitions.

THE UNIVERSITY OF PENNSYLVANIA

[4] E. Hille proved that $f(n) > n^{\rho-\epsilon}$ for infinitely many n (ibid).

ANNALS OF MATHEMATICS
Vol. 42, No. 4, October, 1941

CONCRETE REPRESENTATION OF ABSTRACT (M)-SPACES
(A characterization of the space of continuous functions)

BY SHIZUO KAKUTANI

(Received December 23, 1940)

Introduction

A Banach space is called an abstract (M)-space if it is a linear lattice, and if it satisfies, besides the usual axioms of Banach lattices, the following condition:

$$x \geqq 0, \qquad y \geqq 0 \qquad \textit{imply } \| x \vee y \| = \max (\| x \|, \| y \|).$$

The principal purpose of the present paper is to show that every such abstract (M)-space can be concretely represented isometrically and lattice isomorphically by a subspace of the space $C(\Omega)$ of all continuous real valued functions defined on a certain compact Hausdorff space Ω.[1] This may be viewed as a characterization of the subspace of the space of continuous functions, which is closed both in a sense of topology and lattice. Furthermore, this problem has close relations with the problems of concrete representation of abstract (L)-spaces, discussed by the author [9], [11] in connection with the general theory of Markoff processes and ergodic theorems.

In Part I, we shall give the definitions and the exact formulations of fundamental theorems (Theorems 1 and 2). We shall also discuss their relations to other problems concerning the concrete representation of abstract spaces and rings. The proofs of Theorems 1 and 2 will be given in Part II. The main idea is to consider a class of functionals (viz., linear- and lattice-homomorphisms of a given space to the space of real numbers) and to use the notion of weak topology. The last Part III is devoted to the applications of fundamental theorems, and the relations with the theory of abstract (L)-spaces will also be discussed.

The principal results and the general outline of their proofs were previously announced in [10]. I should like to express my hearty thanks to Professors H. F. Bohnenblust, C. Chevalley and J. von Neumann for their kind discussions in the course of this work.[2]

---

[1] We use the term "compact" as "open coverings can be reduced to finite coverings."

[2] After completing the paper, I noticed that the same problem was discussed by Mark and Selim Krein (On an inner characteristic of the set of all continuous functions defined on a bicompact Hausdorff space, C. R. URSS, 27 (1940), 427–430). In this paper the existence of a unit element is always assumed; Theorem 2 was proved by them, but the case of Theorem 1 was not discussed.

I. Definitions and the Statement of Fundamental Theorems

1. Definition of abstract (M)-spaces

A Banach space (AM) is called an *abstract* (M)*-space*, if there is defined a relation $x \geqq y$ (or equivalently $y \leqq x$) for some pairs of elements $x, y \in (AM)$, and if it satisfies the following conditions ($x, y, z, w \in (AM)$, λ: real number):

(1.1) $x \geqq y, \quad y \geqq x$ *imply* $x = y,$

(1.2) $x \geqq y, \quad y \geqq z$ *imply* $x \geqq z,$

(1.3) $x \geqq y, \quad \lambda \geqq 0$ *imply* $\lambda x \geqq \lambda y,$

(1.4) $x \geqq y$ *implies* $x + z \geqq y + z,$

(1.5) *To any pair of elements* $x, y \in (AM)$, *there exists a maximum* $z = x \vee y$ *such that* $z \geqq x, z \geqq y$ *and* $z' \geqq z$ *for any* z' *with* $z' \geqq x, z' \geqq y.$

(1.6) *To any pair of elements* $x, y \in (AM)$, *there exists a minimum* $w = x \wedge y$ *such that* $w \leqq x, w \leqq y$ *and* $w' \leqq w$ *for any* w' *with* $w' \leqq x, w' \leqq y.$

(1.7) $x_n \geqq y_n, x_n \to x$ (*strongly*), $y_n \to y$ (*strongly*) *imply* $x \geqq y,$

(1.8) $x \wedge y = 0$ *implies* $\| x + y \| = \| x - y \|,$

(1.9) $x \geqq 0, \quad y \geqq 0$ *imply* $\| x \vee y \| = \max (\| x \|, \| y \|).$

The conditions (1.1)–(1.6) mean that (AM) is a linear lattice (or a vector lattice). Such a linear space was discussed by G. Birkhoff [1], H. Freudenthal [5] and L. Kantorovitch [12]. The general theory of such linear lattices will be found in G. Birkhoff [2]. Here we shall state some of the fundamental properties of such spaces, which we shall need in the following discussions:

(1.10) $\lambda \geqq 0$ *implies* $\lambda(x \vee y) = \lambda x \vee \lambda y,$
$\lambda(x \wedge y) = \lambda x \wedge \lambda y,$

(1.11) $(x \vee y) + z = (x + z) \vee (y + z),$
$(x \wedge y) + z = (x + z) \wedge (y + z),$

(1.12) $(x \vee y) + (x \wedge y) = x + y,$

(1.13) $x = x \vee 0 - (-x) \vee 0, \quad (x \vee 0) \wedge ((-x) \vee 0) = 0,$

(1.14) $(x \vee y) \wedge z = (x \wedge z) \vee (y \wedge z),$

(1.15) $(x \wedge y) \vee z = (x \vee z) \wedge (y \vee z).$

We shall make use of these properties without any reference to them.

Among all conditions (1.1)–(1.9), the last condition (1.9) is the most important, although, as is easily seen, the condition (1.8) is also indispensable in the following discussions.

It is clear that (1.9) implies:

(1.16) $x \geqq y \geqq 0$ *implies* $\| x \| \geqq \| y \| \geqq 0$,

(1.17) $x \wedge y = 0$ *implies* $\| x \vee y \| = \max(\| x \|, \| y \|)$.

These conditions were introduced by H. F. Bohnenblust [3]. But we do not know whether or not, conversely, the conditions (1.16) and (1.17) together imply (1.9).[2,a]

2. Examples of abstract (M)-spaces

The foremost example of abstract (M)-spaces will be given by the space $C(\Omega)$ of all bounded continuous real-valued functions $x(t)$ defined on a Hausdorff space Ω where $\| x \| = \text{l.u.b.}_{t \in \Omega} | x(t) |$ and $x \geqq y$ if and only if $x(t) \geqq y(t)$ for any $t \in \Omega$. If we now take an arbitrary set Ω and consider it as a discrete space, then the space $C(\Omega)$ is nothing but the space $M(\Omega)$ of all bounded real-valued functions $x(t)$ defined on Ω (with the same norm and partial ordering as in the case of $C(\Omega)$).

Moreover, if we consider a subspace $M(\Omega; m)$ of $M(\Omega)$ consisting of all bounded measurable real-valued functions $x(t)$ defined on Ω (where measurability is with respect to a measure $m(E)$ defined *a priori* on Ω), then this is also an example of abstract (M)-spaces. Furthermore, a space of the same kind will be obtained if we neglect some class of subsets of Ω. For example, in the case of $M(\Omega; m)$, if we neglect sets of measure zero of Ω, then the space $M'(\Omega; m)$ thus obtained is again an example of abstract (M)-spaces. It is clear that in this case we have to put $\| x \| = \text{ess. max.}_{t \in \Omega} | x(t) |$[3] and $x \geqq y$ if and only if $x(t) \geqq y(t)$ almost everywhere on Ω. (Two functions which differ from each other only on a set of measure zero are considered as the same element of $M'(\Omega; m)$.)

Another important example of abstract (M)-spaces will be obtained by considering a special subspace of $C(\Omega)$. Indeed, if we consider only those functions $x(t)$ of $C(\Omega)$ which vanish at a given point $t_0 \in \Omega$, then the space $C(\Omega; t_0)$ of all such functions $x(t)$ will be an example. More generally, if we consider the subspace $C(\Omega; t_0, t_0'; \lambda_0)$ of $C(\Omega)$ consisting of all functions $x(t) \in C(\Omega)$ which satisfy the relation:

$$x(t_0) = \lambda_0 x(t_0'), \tag{2.1}$$

where $t_0, t_0' \in \Omega$ and $0 \leqq \lambda_0 \leqq 1$, then this is again an example. Furthermore, it will be easily seen that the spaces of the same kind will be obtained if we replace the point t_0 of the former example by a closed set F, and if we replace the relation (2.1) in the latter example by the system of relations:

$$x(t_\alpha) = \lambda_\alpha x(t_\alpha'), \qquad \alpha \in \mathfrak{M}, \tag{2.2}$$

[2a] See the following paper.

[3] ess. max.$_{t \in \Omega} | x(t) |$ means the greatest lower bound of all $\alpha > 0$ such that $m[t: | x(t) | > \alpha] = 0$.

where $t_\alpha, t'_\alpha \in \Omega$, $0 \leqq \lambda_\alpha \leqq 1$ and $\mathfrak{M}$ is a set of indices α of any power. We shall denote these spaces by $C(\Omega; F)$ and $C(\Omega; t_\alpha, t'_\alpha; \lambda_\alpha; \alpha \in \mathfrak{M})$ respectively.

It is to be noted that, if in the case of $C(\Omega; F)$, the closed set F consists of more than one point, then we can identify all points of F, and that we can thus reduce the space $C(\Omega; F)$ to the case of $C(\Omega; t_0)$. This fact is also true for the spaces $C(\Omega; t_\alpha, t'_\alpha; \lambda_\alpha; \alpha \in \mathfrak{M})$ if there exists a λ_α with $\lambda_\alpha = 1$.

3. Unit element

If there exists an element $\mathbf{1}$ such that

$$\mathbf{1} \geqq 0, \qquad \|\mathbf{1}\| = 1, \tag{3.1}$$

$$\|x\| \leqq 1 \quad \textit{implies} \quad x \leqq \mathbf{1}, \tag{3.2}$$

then $\mathbf{1}$ is called a *unit element*.

It is clear that there exists at most one unit element. But the existence of a unit element is not assumed in abstract (M)-spaces. Indeed, among the examples given in section 2, the spaces $C(\Omega)$, $M(\Omega)$, $M(\Omega; m)$ and $M'(\Omega; m)$ have unit elements (in this case, $x(t) \equiv 1$ is a unit element), while this is not the case for $C(\Omega; t_0)$ and $C(\Omega; t_0, t'_0; \lambda_0)$ if the point t_0 is not isolated in Ω.

REMARK 1. The existence of a unit element together with the condition (1.16) will imply (1.9). Indeed, given $x \geqq 0$, $y \geqq 0$, we have, by (3.2), $x \leqq \|x\| \cdot \mathbf{1}$, $y \leqq \|y\| \cdot \mathbf{1}$, and consequently $x \vee y \leqq \max(\|x\|, \|y\|) \cdot \mathbf{1}$, which in turn implies $\|x \vee y\| \leqq \max(\|x\|, \|y\|)$ by (3.1). The inverse relation $\|x \vee y\| \geqq \max(\|x\|, \|y\|)$ is clear from (1.16).

REMARK 2. It will be easily seen that $\mathbf{1}$ is also a unit element in the sense of H. Freudenthal [5] (notation: F-unit) which is defined by the following condition:

$$x > 0 \quad \textit{implies} \quad x \wedge \mathbf{1} > 0. \tag{3.3}$$

Indeed, in case $\|x\| \leqq 1$ this follows directly from (3.2), and if $\|x\| > 1$, then $x \wedge \mathbf{1} \geqq \|x\|^{-1} \cdot x \wedge \mathbf{1} = \|x\|^{-1} \cdot x > 0$.

Moreover, an F-unit can exist even if the unit element in our sense does not exist. For example, consider the space $C(\Omega_0; 0)$ or $C(\Omega_0; 0, 1; \frac{1}{2})$, where Ω_0 is the closed interval $0 \leqq t \leqq 1$. These are the spaces of all continuous real-valued functions $x(t)$ defined on $0 \leqq t \leqq 1$ such that $x(0) = 0$ or $x(0) = \frac{1}{2}x(1)$ respectively. It is easy to see that there exists an F-unit in these spaces, while there is no unit in our sense at all. This fact is also true for general spaces $C(\Omega; t_0)$ and $C(\Omega; t_0, t'_0; \lambda_0)$ $(\lambda_0 < 1)$ if the point t_0 in question is not isolated in Ω, and if the topology of Ω satisfies the first countability axiom at $t = t_0$. But in a more general case when the topology of Ω does not satisfy the first countability axiom at $t = t_0$, this is no longer true. Indeed, let us consider the space $D(\Omega)$ of all bounded real-valued functions $x(t)$ defined on Ω such that the set of points at which we have $|x(t)| > \epsilon$ is finite for any $\epsilon > 0$. Then $D(\Omega)$ is an abstract (M)-space if we define the norm and the partial ordering in

the same way as in the case of $C(\Omega)$. It is easy to see that there exists no F-unit in this (M)-space, and it is not without interest to notice that, if we adjoin a new point t_0 to Ω and if we introduce a topology of $\Omega^* = \Omega + (t_0)$ in the following manner:

(3.4) *Every point of Ω is isolated,*

(3.5) *Every neighborhood $U(t_0)$ of t_0 is a set of the form: $U(t_0) = (t_0) + (\Omega - A)$, where A is an arbitrary finite subset of Ω,* then Ω^* is a compact Hausdorff space and the space $C(\Omega; t_0)$ is the space $D(\Omega)$ in question.

4. Fundamental theorems

After the preliminaries given above, we shall now state the fundamental theorems.

THEOREM 1. *For any abstract (M)-space (AM) (with or without a unit element), there exists a compact Hausdorff space Ω and a system of pairs of points $\{t_\alpha\}$, $\{t'_\alpha\}$ $(t_\alpha \in \Omega, \alpha \in \mathfrak{M})$ and real numbers $\{\lambda_\alpha\}$ $(0 \leqq \lambda_\alpha < 1)$ $(\alpha \in \mathfrak{M})$ such that (AM) is isometric and lattice isomorphic to the space $C(\Omega; t_\alpha, t'_\alpha; \lambda_\alpha; \alpha \in \mathfrak{M})$ of all bounded continuous real-valued functions $x(t)$ which are defined on Ω and which satisfy the following relations:*

$$x(t_\alpha) = \lambda_\alpha x(t'_\alpha), \qquad \alpha \in \mathfrak{M}, \tag{4.1}$$

where $\mathfrak{M}$ is a set of indices α whose power can be arbitrarily large.

This theorem is complicated, and the reason for this is that we do not assume the existence of a unit element. If we assume the existence of a unit element, then it becomes simpler:

THEOREM 2. *For any abstract (M)-space (AM) with a unit (see section 3), there exists a compact Hausdorff space Ω such that (AM) is isometric and lattice isomorphic to the space $C(\Omega)$ of all bounded continuous real-valued functions $x(t)$ defined on Ω.*[4]

Proofs of Theorems 1 and 2 will be given in sections 7 and 11 respectively. We can also prove that the compact Hausdorff space Ω is uniquely determined up to a homeomorphism. This problem will be discussed in section 13 (Theorem 6).

5. Relations to other problems

As is easily observed, this problem has close relations to the problems discussed by I. Gelfand [6], I. Gelfand–A. Kolmogoroff [7], J. von Neumann [13], G. Šilov [16], M. H. Stone [17], [18], and K. Yosida [19]. In all cases discussed by these authors, the Banach space in question was always considered as a ring, and they have obtained a concrete representation of such Banach rings. In our case, however, the space in question is a lattice instead of being a ring, and we can also obtain a concrete representation of such Banach lattices.

Further, it is to be noted that the existence of the unit element (which is

[4] It is not difficult to see that Ω satisfies the second countability axiom of Hausdorff (and hence metrisable) if and only if the given (AM) is separable. It is also obvious that Ω consists only of a finite number of points if and only if (AM) is of finite dimension.

necessary in the case of a ring) is not assumed in Theorem 1. H. F. Bohnenblust [3] has recently discussed the characterization of L_p spaces ($1 \leqq p \leqq \infty$), and the case $p = \infty$ corresponds to our problem. He has only assumed the conditions (1.16) and (1.17), and has obtained a concrete representation of such Banach lattices. But the existence of an F-unit is assumed in his paper, and moreover, besides the separability of the space in question, he assumed that

(5.1) $0 \leqq x_1 \leqq x_2 \leqq \cdots \leqq x_n \leqq \cdots \leqq \bar{x}$ *implies the existence of* $\sup_n x_n$,

where $x' = \sup_n x_n$ is an element such that $x_n \leqq x'$ for $n = 1, 2, \cdots$, and $x' \leqq x''$ for any x'' with $x_n \leqq x''$ for $n = 1, 2, \cdots$. This condition (5.1), which is also assumed by J. v. Neumann [13] and M. H. Stone [18] (these authors introduced the notions of both ring and lattice), is not satisfied even in the case of the space $C(\Omega_0)$ of all continuous real-valued functions $x(t)$ defined on $\Omega_0 : 0 \leqq t \leqq 1$. Indeed, for example, the sequence of functions $x_n(t)$ $(n = 1, 2, \cdots)$, where

$$x_n(t)\begin{cases} = 0, & 0 \leqq t \leqq \frac{1}{2}, \\ = n(t - \frac{1}{2}), & \frac{1}{2} \leqq t \leqq \frac{1}{2} + \frac{1}{n}, \\ = 1, & \frac{1}{2} + \frac{1}{n} \leqq t \leqq 1, \end{cases}$$

satisfies (5.1) with $\bar{x}(t) \equiv 1$, while there exists no such continuous function $x'(t)$ which corresponds to the element $x' = \sup_n x_n$.

Finally, as will be seen from the papers of J. v. Neumann [13], M. H. Stone [18] and K. Yosida [19], our problem has also a close connection with the theory of spectra for operators.[5]

---

[5] As was remarked by Professor Stone, our problem has close connections with the theory of harmonic functions. Indeed, let D be an arbitrary domain in the Gaussian plane (we do not assume that D is simply connected), and let $H(D)$ be the space of all real valued bounded harmonic functions $u(z)$ defined in D. $H(D)$ is an abstract (M)-space with respect to the norm $\| u \| = \sup_{z \in D} | u(z) |$ and the partial ordering $u \geqq v$: $u(z) \geqq v(z)$ for all $z \in D$. It is to be noticed that $w = u \vee v$ is not equal to the max. $(u(z), v(z))$ in the ordinary sense (maximum taken at each point of D), but $w(z)$ is the smallest harmonic function which satisfies $w(z) \geqq u(z)$ and $w(z) \geqq v(z)$ for all $z \in D$. (The existence of such $w(z)$ follows from the well known principle in the theory of harmonic functions). If, in particular, D is the interior of the unit circle, then every $u(z) \in H(D)$ can be expressed by the Poisson integral:

$$(*) \qquad u(z) = u(re^{i\theta}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{(1 - r^2)x(t)\,dt}{1 - 2r\cos(\theta - t) + r^2},$$

where $x(t)$ is a bounded measurable function defined on the periphery Γ of the unit circle. It is not difficult to see that, if we consider the space $M'(\Gamma, m)$ of all real valued bounded measurable functions $x(t)$ defined on Γ (with respect to the ordinary Lebesgue measure m on Γ; we consider two functions which differ on a set of measure zero to be the same element of $M'(\Gamma, m)$; see section 2), then $M'(\Gamma, m)$ is isometric and lattice isomorphic to the abstract (M)-space $H(D)$, and this isomorphism is given by (*). It is to be noticed that in this case both $H(D)$ and $M'(\Gamma, m)$ satisfy the condition (5.1), and that the representation (*) is isomorphic even with respect to the operation $\sup_n x_n$.

II. Proofs of the Fundamental Theorems

6. Weak topology of Banach spaces

Before going to the proofs of the fundamental theorems, we shall first enter into some elementary considerations on the weak topology of Banach spaces.

Let E be a Banach space and let E^* be its conjugate space, i.e., the Banach space of all bounded linear functionals $f(x)$ defined on E with $\|f\| = \text{l.u.b.}_{\|x\|\leq 1}|f(x)|$ as its norm.

We shall introduce a weak topology[6] into E^*. For any $f_0 \in E^*$ its weak neighborhood $U(f_0 ; x_1 , x_2 , \cdots , x_n ; \epsilon)$ is defined as the totality of all $f \in E^*$ such that $|f(x_i) - f_0(x_i)| < \epsilon$ for $i = 1, 2, \cdots , n$, where $\{x_i\}$ $(i = 1, 2, \cdots , n)$ is an arbitrary system of points from E and $\epsilon > 0$ is an arbitrary positive number. The system of all such weak neighborhoods $U(f_0 ; x_1 , x_2 , \cdots , x_n ; \epsilon)$ determines a Hausdorff topology on E^*, and this topology is called the *weak topology* of E^* (as functionals).

We shall list here some fundamental lemmas on weak topology, which we shall need in the following discussions:[7]

Lemma 6.1. *For any fixed $x_0 \in E$, $x_0(f) = f(x_0)$ is a continuous function (even a bounded linear functional) defined on E^*, with respect to the weak topology of E^*.*

Lemma 6.2. *The unit sphere: $\|f\| \leqq 1$ of E^* is compact with respect to the weak topology of E^*.*

Lemma 6.3. *Let $\Phi(\alpha_1 , \alpha_2 , \cdots , \alpha_n)$ be a continuous function of n variables $\alpha_1 , \alpha_2 , \cdots , \alpha_n$, and let $x_1 , x_2 , \cdots , x_n$ be a given system of points of a Banach space E. Then the totality of all $f \in E^*$ such that $\Phi(f(x_1), f(x_2), \cdots , f(x_n)) = 0$ is a closed subset of E^*.*

7. Outline of the proof of Theorem 1

Let us consider a (real-valued) bounded linear functional $f(x)$ defined on (AM) which satisfies the following conditions:

---

[6] Weak topology was first introduced by J. von Neumann, Zur Algebra der Funktionaloperatoren und Theorie der normalen Operatoren, Math. Ann. 102(1930), 370–427. See also: L. Alaoglu, Weak topologies of normed linear spaces, Annals of Math., 41(1940), 252–267; N. Bourbaki, Sur les espaces de Banach, C. R. Paris, 206(1938), 1701–1704; S. Kakutani, Weak topology and the regularity of Banach spaces, Proc. Imp. Acad. Japan, 15(1939), 169–173; A. E. Taylor, Weak topologies of Banach spaces, Proc. Nat. Acad. Sci. U. S. A., 25(1939), 438–440; A. Tychonoff, Über die topologische Erweiterungen von Räumen, Math. Ann., 102(1930), 544–561; A. Tychonoff, Über einen Funktionenraum, Math. Ann., 111(1935,) 762–766.

[7] Lemmas 6.1 and 6.3 follow directly from the definition of weak topology. Lemma 6.2 is a consequence of Tychonoff's theorem that a topological product of an arbitrary (finite or infinite, not necessarily countable) number of closed intervals is compact. (See the papers of A. Tychonoff cited in footnote 6). A simpler proof of this theorem may be found in the following papers, where the result is extended to the case of the topological product of an arbitrary number of compact spaces: J. W. Alexander, Ordered sets, complexes and the problem of compactification, Proc. Nat. Acad. Sci. U. S. A., 25(1939), 296–298; H. Wallman, Lattices and topological spaces, Annals of Math., 39(1938), 112–126.. See also E. Čech [4].

$$\|f\| = 1, \tag{7.1}$$

$$f(x + y) = f(x) + f(y), \tag{7.2}$$

$$f(\lambda x) = \lambda f(x), \qquad \lambda \gtreqless 0, \tag{7.3}$$

$$x \geqq 0 \quad \textit{implies} \quad f(x) \geqq 0, \tag{7.4}$$

$$x \wedge y = 0 \quad \textit{implies} \quad f(x) = 0 \quad \text{or} \quad f(y) = 0, \tag{7.5}$$

$$f(x \vee y) = \max(f(x), f(y)), \tag{7.6}$$

$$f(x \wedge y) = \min(f(x), f(y)). \tag{7.7}$$

These conditions are clearly not mutually independent. But it is not necessary to discuss their independence. We shall only notice that the conditions (7.6) and (7.7) follow directly from the other conditions. This fact is needed in the following.

If we consider the space of real numbers as an abstract (M)-space (this is indeed the simplest non-trivial example of abstract (M)-spaces), then these conditions mean that $x \to f(x)$ is a continuous homomorphic mapping of a given space (AM) onto the space of real numbers. (For the general principle of the proofs of such theorems see [10].) Indeed, (7.1) implies that $x \to f(x)$ is continuous; (7.2) and (7.3) mean that $x \to f(x)$ is a linear homomorphism; and lastly, (7.4), (7.5), (7.6) and (7.7) mean that $x \to f(x)$ is a lattice homomorphism.

Let us now consider the totality Ω' of all bounded linear functionals $f(x)$ defined on (AM) which satisfy the conditions (7.1)–(7.7). Ω' is a bounded set contained in the unit sphere of the conjugate space $(AM)^*$ of (AM). Consequently, by Lemma 6.2, the closure $\Omega = \overline{\Omega'}$ of Ω' with respect to the weak topology of $(AM)^*$ is compact, and it is easy to see that every $f \in \Omega$ satisfies the condition (7.2)–(7.7). This is a direct consequence of Lemma 6.3. But, in general, the condition (7.1) is not necessarily satisfied, and we can only say that we have $\|f\| \leqq 1$ for any $f \in \Omega$. If there exists, indeed, an element $f^1 \in \Omega$ with $\|f\| = \lambda < 1$, then we have $f = \lambda f'$ for some $f' \in \Omega'$. Let us denote all such relations which exist among the elements of Ω by:

$$f_\alpha = \lambda_\alpha f'_\alpha, \qquad \alpha \in \mathfrak{M}, \tag{7.8}$$

where $f_\alpha \in \Omega$, $f'_\alpha \in \Omega'$, $0 \leqq \lambda_\alpha < 1$ and $\mathfrak{M}$ is a set of indices α which can be finite, denumerably infinite or even non-denumerable.

We shall next consider the functional $x(f) = f(x)$ defined on Ω (x being fixed). For each $x_0 \in (AM)$, by Lemma 6.1, $x_0(f)$ is a bounded continuous real-valued function defined on Ω, and it is clear that the conditions:

$$x(f_\alpha) = \lambda_\alpha x(f'_\alpha), \qquad \alpha \in \mathfrak{M}, \tag{7.9}$$

are all satisfied. Consequently, for each $x_0 \in (AM)$, the function $x_0(f)$ can be considered as an element of $C(\Omega; f_\alpha, f'_\alpha; \lambda_\alpha; \alpha \in \mathfrak{M})$ and the conditions (7.2)–(7.7) imply that $x \to x(f)$ is a continuous (linear and lattice) homomorphism of (AM) into the space $C(\Omega; f_\alpha, f'_\alpha; \lambda_\alpha; \alpha \in \mathfrak{M})$.

II. Proofs of the Fundamental Theorems

6. Weak topology of Banach spaces

Before going to the proofs of the fundamental theorems, we shall first enter into some elementary considerations on the weak topology of Banach spaces.

Let E be a Banach space and let E^* be its conjugate space, i.e., the Banach space of all bounded linear functionals $f(x)$ defined on E with $\|f\| = \text{l.u.b.}_{\|x\|\leqq 1}|f(x)|$ as its norm.

We shall introduce a weak topology[6] into E^*. For any $f_0 \in E^*$ its weak neighborhood $U(f_0 ; x_1 , x_2 , \cdots , x_n ; \epsilon)$ is defined as the totality of all $f \in E^*$ such that $|f(x_i) - f_0(x_i)| < \epsilon$ for $i = 1, 2, \cdots , n$, where $\{x_i\}$ $(i = 1, 2, \cdots , n)$ is an arbitrary system of points from E and $\epsilon > 0$ is an arbitrary positive number. The system of all such weak neighborhoods $U(f_0 ; x_1 , x_2 , \cdots , x_n ; \epsilon)$ determines a Hausdorff topology on E^*, and this topology is called the *weak topology* of E^* (as functionals).

We shall list here some fundamental lemmas on weak topology, which we shall need in the following discussions:[7]

Lemma 6.1. *For any fixed $x_0 \in E$, $x_0(f) = f(x_0)$ is a continuous function (even a bounded linear functional) defined on E^*, with respect to the weak topology of E^*.*

Lemma 6.2. *The unit sphere: $\|f\| \leqq 1$ of E^* is compact with respect to the weak topology of E^*.*

Lemma 6.3. *Let $\Phi(\alpha_1 , \alpha_2 , \cdots , \alpha_n)$ be a continuous function of n variables $\alpha_1 , \alpha_2 , \cdots , \alpha_n$, and let $x_1 , x_2 , \cdots , x_n$ be a given system of points of a Banach space E. Then the totality of all $f \in E^*$ such that $\Phi(f(x_1), f(x_2), \cdots , f(x_n)) = 0$ is a closed subset of E^*.*

7. Outline of the proof of Theorem 1

Let us consider a (real-valued) bounded linear functional $f(x)$ defined on (AM) which satisfies the following conditions:

---

[6] Weak topology was first introduced by J. von Neumann, Zur Algebra der Funktionaloperatoren und Theorie der normalen Operatoren, Math. Ann. 102(1930), 370–427. See also: L. Alaoglu, Weak topologies of normed linear spaces, Annals of Math., 41(1940), 252–267; N. Bourbaki, Sur les espaces de Banach, C. R. Paris, 206(1938), 1701–1704; S. Kakutani, Weak topology and the regularity of Banach spaces, Proc. Imp. Acad. Japan, 15(1939), 169–173; A. E. Taylor, Weak topologies of Banach spaces, Proc. Nat. Acad. Sci. U. S. A., 25(1939), 438–440; A. Tychonoff, Über die topologische Erweiterungen von Räumen, Math. Ann., 102(1930), 544–561; A. Tychonoff, Über einen Funktionenraum, Math. Ann., 111(1935,) 762–766.

[7] Lemmas 6.1 and 6.3 follow directly from the definition of weak topology. Lemma 6.2 is a consequence of Tychonoff's theorem that a topological product of an arbitrary (finite or infinite, not necessarily countable) number of closed intervals is compact. (See the papers of A. Tychonoff cited in footnote 6). A simpler proof of this theorem may be found in the following papers, where the result is extended to the case of the topological product of an arbitrary number of compact spaces: J. W. Alexander, Ordered sets, complexes and the problem of compactification, Proc. Nat. Acad. Sci. U. S. A., 25(1939), 296–298; H. Wallman, Lattices and topological spaces, Annals of Math., 39(1938), 112–126.. See also E. Čech [4].

$$\|f\| = 1, \tag{7.1}$$

$$f(x + y) = f(x) + f(y), \tag{7.2}$$

$$f(\lambda x) = \lambda f(x), \qquad \lambda \gtreqless 0, \tag{7.3}$$

$$x \geqq 0 \quad \textit{implies} \quad f(x) \geqq 0, \tag{7.4}$$

$$x \wedge y = 0 \quad \textit{implies} \quad f(x) = 0 \quad \text{or} \quad f(y) = 0, \tag{7.5}$$

$$f(x \vee y) = \max(f(x), f(y)), \tag{7.6}$$

$$f(x \wedge y) = \min(f(x), f(y)). \tag{7.7}$$

These conditions are clearly not mutually independent. But it is not necessary to discuss their independence. We shall only notice that the conditions (7.6) and (7.7) follow directly from the other conditions. This fact is needed in the following.

If we consider the space of real numbers as an abstract (M)-space (this is indeed the simplest non-trivial example of abstract (M)-spaces), then these conditions mean that $x \to f(x)$ is a continuous homomorphic mapping of a given space (AM) onto the space of real numbers. (For the general principle of the proofs of such theorems see [10].) Indeed, (7.1) implies that $x \to f(x)$ is continuous; (7.2) and (7.3) mean that $x \to f(x)$ is a linear homomorphism; and lastly, (7.4), (7.5), (7.6) and (7.7) mean that $x \to f(x)$ is a lattice homomorphism.

Let us now consider the totality Ω' of all bounded linear functionals $f(x)$ defined on (AM) which satisfy the conditions (7.1)–(7.7). Ω' is a bounded set contained in the unit sphere of the conjugate space $(AM)^*$ of (AM). Consequently, by Lemma 6.2, the closure $\Omega = \overline{\Omega'}$ of Ω' with respect to the weak topology of $(AM)^*$ is compact, and it is easy to see that every $f \in \Omega$ satisfies the condition (7.2)–(7.7). This is a direct consequence of Lemma 6.3. But, in general, the condition (7.1) is not necessarily satisfied, and we can only say that we have $\|f\| \leqq 1$ for any $f \in \Omega$. If there exists, indeed, an element $f^\lambda \in \Omega$ with $\|f\| = \lambda < 1$, then we have $f = \lambda f'$ for some $f' \in \Omega'$. Let us denote all such relations which exist among the elements of Ω by:

$$f_\alpha = \lambda_\alpha f'_\alpha, \qquad \alpha \in \mathfrak{M}, \tag{7.8}$$

where $f_\alpha \in \Omega$, $f'_\alpha \in \Omega'$, $0 \leqq \lambda_\alpha < 1$ and $\mathfrak{M}$ is a set of indices α which can be finite, denumerably infinite or even non-denumerable.

We shall next consider the functional $x(f) = f(x)$ defined on Ω (x being fixed). For each $x_0 \in (AM)$, by Lemma 6.1, $x_0(f)$ is a bounded continuous real-valued function defined on Ω, and it is clear that the conditions:

$$x(f_\alpha) = \lambda_\alpha x(f'_\alpha), \qquad \alpha \in \mathfrak{M}, \tag{7.9}$$

are all satisfied. Consequently, for each $x_0 \in (AM)$, the function $x_0(f)$ can be considered as an element of $C(\Omega; f_\alpha, f'_\alpha; \lambda_\alpha; \alpha \in \mathfrak{M})$ and the conditions (7.2)–(7.7) imply that $x \to x(f)$ is a continuous (linear and lattice) homomorphism of (AM) into the space $C(\Omega; f_\alpha, f'_\alpha; \lambda_\alpha; \alpha \in \mathfrak{M})$.

Thus, all that we have to do in order to prove Theorem 1, is to show that this mapping $x \to x(f)$ is an isometric mapping of (AM) onto $C(\Omega; f_\alpha, f'_\alpha; \lambda_\alpha; \alpha \in \mathfrak{M})$. It will be easily seen that we have, for this purpose, only to prove the following lemmas:

LEMMA 7.1. *For any $x_0 \in (AM)$, there exists an $f_0 \in \Omega$ such that $|f_0(x_0)| = \|x_0\|$.*

LEMMA 7.2. *For any continuous real-valued function $X_0(f) \in C(\Omega; f_\alpha, f'_\alpha; \lambda_\alpha; \alpha \in \mathfrak{M})$, there exists an $x_0 \in (AM)$ such that $f(x_0) = X_0(f)$ for any $f \in \Omega$.*

These are the fundamental lemmas, to whose proofs the remaining part of Part II is devoted. The final proofs of these lemmas will be given in sections 9 and 10 respectively.

8. Lemmas on abstract (M)-spaces

Let $x_0 \in (AM)$ and let $f(x)$ be a bounded linear functional defined on (AM).

LEMMA 8.1.[8] *$x_0 \geqq 0$, $\|x_0\| = 1$, $f(x_0) = 1$, $\|f\| = 1$ imply $f(y) \geqq 0$ for any $y \geqq 0$; i.e., $f(x)$ is a positive linear functional.*

PROOF. Without loss of generality, we may assume $\|y\| = 1$. Then, $0 \leqq (x_0 - y) \vee 0 \leqq x_0$ implies $\|(x_0 - y) \vee 0\| \leqq \|x_0\| = 1$, and $0 \leqq (y - x_0) \vee 0 \leqq y$ implies $\|(y - x_0) \vee 0\| \leqq \|y\| = 1$. Hence, $\|x_0 - y\| = \|(x_0 - y) \vee 0 - (y - x_0) \vee 0\| = \|(x_0 - y) \vee 0 + (y - x_0) \vee 0\| = \max(\|(x_0 - y) \vee 0\|, \|(y - x_0) \vee 0\|) \leqq 1$, and consequently $1 - f(y) = f(x_0) - f(y) = f(x_0 - y) \leqq \|f\| \cdot \|x_0 - y\| \leqq 1$, or equivalently $f(y) \geqq 0$.

LEMMA 8.2. *$x_0 \geqq 0$, $\|x_0\| = 1$, $f(x_0) = 1$, $\|f\| = 1$, $y \geqq 0$, $\|x_0 \wedge y\| < 1$, imply $f(y) = f(x_0 \wedge y) < 1$. In particular, $x_0 \wedge y = 0$ implies $f(y) = 0$.*

PROOF. Put $z = (1 - \lambda)(x_0 \wedge y) + \lambda y$, where $0 < \lambda < 1$. Then $x_0 \wedge y \leqq z \leqq y$ and, for sufficiently small λ, we have $\|z\| \leqq (1 - \lambda)\|x_0 \wedge y\| + \lambda\|y\| < 1$. Hence, $x_0 \wedge z = x_0 \wedge (y \wedge z) = (x_0 \wedge y) \wedge z = x_0 \wedge y$, and, since by lemma 8.1 $f(x)$ is a positive linear functional, $1 = f(x_0) \leqq f(x_0 \vee z) \leqq \|f\| \cdot \|x_0 \vee z\| = \|f\| \cdot \max(\|x_0\|, \|z\|) \leqq 1$, or $f(x_0 \vee z) = 1$. Consequently, $f(x_0 \wedge y) = f(x_0 \wedge z) = f(x_0 + z - x_0 \vee z) = f(x_0) + f(z) - f(x_0 \vee z) = 1 + f(z) - 1 = f(z)$, i.e., $f(x_0 \wedge y) = f(z) = (1 - \lambda)f(x_0 \wedge y) + \lambda f(y)$. Since $\lambda > 0$, we must have $f(x_0 \wedge y) = f(y)$.

LEMMA 8.3. *For any $x_0 \geqq 0$ with $\|x_0\| = 1$, and for any system $\Sigma = \{x_1, x_2, \cdots, x_n\}$ of positive elements of (AM), there exists a $y_0 \in (AM)$ and another system $\{y_1, y_2, \cdots, y_n\}$ of positive elements of (AM) such that*

$$0 \leqq y_0 \leqq x_0, \qquad \|y_0\| = 1, \tag{8.1}$$

$$\|y_0 \wedge \lambda y_i\| < 1 \quad \text{for any} \quad \lambda > 0 \quad \text{and for} \quad i = 1, 2, \cdots, n, \tag{8.2}$$

$$x_i \wedge x_j = 0 \quad \text{implies} \quad x_i = y_i \quad \text{or} \quad x_j = y_j. \tag{8.3}$$

PROOF. We shall define z_i and y_i $(i = 1, 2, \cdots, n)$ by mathematical induction. Put $z_0 = x_0$, and let us assume that $z_{i-1} \geqq 0$ with $\|z_{i-1}\| = 1$ and $y_{i-1} \geqq 0$

[8] Lemmas 8.1 and 8.2 are due to Professor Bohnenblust. By virtue of these lemmas, the proof of Lemma 7.1 became simpler. The original proof of the author was complicated.

are already defined. Then, if there exists a $\lambda_i > 0$ such that $\| z_{i-1} \wedge \lambda_i x_i \| = 1$, put $z_i = z_{i-1} \wedge \lambda_i x_i$ and $y_i = 0$. In other cases (i.e., if we have $\| z_{i-1} \wedge \lambda x_i \| < 1$ for any $\lambda > 0$), put $z_i = z_{i-1}$ and $y_i = x_i$. Thus z_i and y_i are determined for $i = 1, 2, \cdots, n$. If we now put $y_0 = z_n$, then this y_0 and the system $\{y_1, y_2, \cdots, y_n\}$ thus obtained are the required ones.

Indeed, (8.1) is clear. (8.2) is also clear, since $y_i > 0$ implies $y_i = x_i$ and consequently $\| y_0 \wedge \lambda y_i \| = \| y_0 \wedge \lambda x_i \| \leqq \| z_{i-1} \wedge \lambda x_i \| < 1$ for any $\lambda > 0$. In order to prove (8.3), assume $x_i \wedge x_j = 0$ and $i < j$. Then $x_i \neq y_i$ implies $z_i = z_{i-1} \wedge \lambda_i x_i$ for some $\lambda_i > 0$, and consequently $0 \leqq z_{j-1} \wedge \lambda x_j \leqq z_i \wedge \lambda x_j = (z_{i-1} \wedge \lambda_i x_i) \wedge \lambda x_j \leqq \lambda_i x_i \wedge \lambda x_j = 0$ for any $\lambda > 0$, which in turn implies $x_j = y_j$.

LEMMA 8.4. *For any $x_0 \geqq 0$ with $\| x_0 \| = 1$ and for any system $\Sigma = \{x_1, x_2, \cdots, x_n\}$ of positive elements of (AM), there exists a bounded linear functional $f_\Sigma(x)$ defined on (AM) such that*

$$f_\Sigma(x_0) = 1, \qquad \| f_\Sigma \| = 1, \tag{8.4}$$

$$f_\Sigma(y) \geqq 0 \quad \textit{for any} \quad y \geqq 0, \tag{8.5}$$

$$x_i \wedge x_j = 0 \quad \textit{implies} \quad f_\Sigma(x_i) = 0 \quad \textit{or} \quad f_\Sigma(x_j) = 0. \tag{8.6}$$

PROOF. Let y_0 and a system $\{y_1, y_2, \cdots, y_n\}$ be defined as in Lemma 8.3. If we consider a bounded linear functional $f_\Sigma(x)$ which satisfies

$$f_\Sigma(y_0) = 1, \qquad \| f_\Sigma \| = 1 \tag{8.7}$$

(the existence of such $f_\Sigma(x)$ is an easy consequence of the Hahn-Banach extension theorem), then this $f_\Sigma(x)$ is the required one. Indeed, (8.5) is clear by Lemma 8.1, and consequently (8.4) becomes also clear since $1 = f_\Sigma(y_0) \leqq f_\Sigma(x_0) \leqq \| f_\Sigma \| \cdot \| x_0 \| = 1$. In order to prove (8.6) we observe that, by Lemma 8.3, we have either $x_i = y_i$ or $x_j = y_j$. Let us assume $x_i = y_i$. Then, by (8.2) we have $\| y_0 \wedge \lambda x_i \| < 1$ for any $\lambda > 0$. Consequently, Lemma 8.2 implies $f_\Sigma(x_i) = \lambda^{-1} f_\Sigma(\lambda x_i) < \lambda^{-1}$, and since $\lambda > 0$ is arbitrary, we must have $f_\Sigma(x_i) = 0$. Since the case $x_j = y_j$ can be treated in the same way, we have proved (8.6), and the proof of Lemma 8.4 is completed.

9. Proof of Lemma 7.1

First we observe that we may assume $x_0 \geqq 0$ and $\| x_0 \| = 1$. This follows easily from the fact that we have $\| x_0 \| = \| x_0 \vee 0 - (-x_0) \vee 0 \| = \| x_0 \vee 0 + (-x_0) \vee 0 \| = \max(\| x_0 \vee 0 \|, \|(-x_0) \vee 0 \|)$. Indeed, it is clear that we may assume $\| x_0 \| = 1$, and in this case we have either $\| x_0 \vee 0 \| = 1$ or $\|(-x_0) \vee 0 \| = 1$. Let us assume $\| x_0 \vee 0 \| = 1$. If we now assume that Lemma 7.1 is proved for $x_0 \geqq 0$ with $\| x_0 \| = 1$, then there exists a bounded linear functional $f_0 \in \Omega$ such that $f_0(x_0 \vee 0) = 1$ and $\| f_0 \| = 1$. Since $(x_0 \vee 0) \wedge ((-x_0) \vee 0) = 0$, we must have $f_0((-x_0) \vee 0) = 0$ by Lemma 8.2, and consequently $f_0(x_0) = f_0(x_0 \vee 0) - f_0((-x_0) \vee 0) = f_0(x_0 \vee 0)$. Since the case $\|(-x_0) \vee 0 \| = 1$ can be treated analogously, we have to prove Lemma 7.1 only for the case $x_0 \geqq 0$, $\| x_0 \| = 1$.

Let now $x_0 \geqq 0$ be a positive element with $\| x_0 \| = 1$. Then, for any system $\Sigma = \{x_1, x_2, \cdots, x_n\}$ of positive elements of (AM), there exists by Lemma 8.4, a bounded linear functional $f_\Sigma(x)$ which satisfies the conditions (8.4), (8.5) and (8.6). Let Ω_Σ be the set of all such functionals $f_\Sigma(x)$. (Ω_Σ is not necessarily a subset of Ω.) Then Ω_Σ is a non-empty compact set with respect to the weak topology of $(AM)^*$. (See Lemmas 6.2 and 6.3.) Moreover, the family $\{\Omega_\Sigma\}$ of all such Ω_Σ for all choice of Σ has a finite intersection property. This follows directly from the fact that we have $\prod_{i=1}^n \Omega_{\Sigma_i} \supseteq \Omega_{\Sigma_1+\Sigma_2+\cdots+\Sigma_n} \neq 0$ for any Σ_i ($i = 1, 2, \cdots, n$). Hence there exists, by a well known argument concerning compact sets, a bounded linear functional $f_0(x)$ which belongs to all Ω_Σ. We shall show that this $f_0(x)$ is the required one. Indeed, it is easy to see that $f_0(x)$ has the properties (7.1)–(7.5) (see Lemmas 8.1, 8.4. (7.1) follows from the fact that $f_0(x) = 1$ and $\| f_0 \| \leqq 1$), and since both (7.6) and (7.7) are the direct consequences of (7.1)–(7.5), the proof of Lemma 7.1 is completed.

10. Proof of Lemma 7.2[9]

Let $X_0(f)$ be an arbitrary bounded continuous real-valued function of $C(\Omega; f_\alpha, f'_\alpha; \lambda_\alpha; \alpha \in \mathfrak{M})$. We shall first prove that for any pair of elements f_0 and g_0 of Ω there exists an element $y_0 \in (AM)$ such that

$$f_0(y_0) = X_0(f_0), \qquad g_0(y_0) = X_0(g_0). \tag{10.1}$$

Indeed, this is clear if there exists an $\alpha \in \mathfrak{M}$ such that $f_0 = f_\alpha$, $g_0 = f'_\alpha$ (or $f_0 = f'_\alpha$, $g_0 = f_\alpha$). In other cases, there exists a pair of points $y'_0, y''_0 \in (AM)$ such that

$$\delta = f_0(y'_0)g_0(y''_0) - f_0(y''_0)g_0(y'_0) \neq 0.$$

Hence, if we put $\alpha = \delta^{-1}(g_0(y''_0)X_0(f_0) - f_0(y''_0)X_0(g_0))$, $\beta = \delta^{-1}(g_0(y'_0)X_0(f_0) - f_0(y'_0)X_0(g_0))$, then (10.1) is clearly satisfied by $y_0 = \alpha y'_0 - \beta y''_0$.

Thus it is proved that there exists, for any $f_0, g_0 \in \Omega$, a point $y_0 \in (AM)$ such that (10.1) is true. Consequently, for any $\epsilon > 0$, there exists a neighborhood $U(f_0)$ of f_0 such that $f(y_0) > X_0(f) - \epsilon$ for any $f \in U(f_0)$. Let us now consider f_0 as a variable ($g_0 \in \Omega$ and $\epsilon > 0$ being fixed). Then there exists, by the compactness of Ω, a finite number of elements $f_i \in \Omega$ ($i = 1, 2, \cdots, m$), their corresponding points $y_i \in (AM)$ ($i = 1, 2, \cdots, m$) and a system of neighborhoods $U(f_i)$ ($i = 1, 2, \cdots, m$) such that $g_0(y_i) = X_0(g_0)$ ($i = 1, 2, \cdots, m$), $\Omega \subseteq \sum_{i=1}^m U(f_i)$ and $f(y_i) > X_0(f) - \epsilon$ for any $f \in U(f_i)$. Consequently, if we put $z_0 = y_1 \vee y_2 \vee \cdots \vee y_m$, then we have $g_0(z_0) = \max_{1 \leq i \leq m} g_0(y_i) = X_0(g_0)$ and $f(z_0) = \max_{1 \leq i \leq m} f(y_i) > X_0(f) - \epsilon$ for any $f \in \Omega$.

Thus we have seen that there exists, for any $g_0 \in \Omega$ and for any $\epsilon > 0$, a point $z_0 \in (AM)$ such that $g_0(z_0) = X_0(g_0)$ and $f(z_0) > X_0(f) - \epsilon$ for any $f \in \Omega$. Clearly, there exists a neighborhood $V(g_0)$ of g_0 such that $f(z_0) < X_0(f) + \epsilon$ for any $f \in V(g_0)$. Let us now consider g_0 as a variable ($\epsilon > 0$ being fixed). Then there

[9] The proof of Lemma 7.2 became simpler by a suggestion of Professor Chevalley.

exist, again by the compactness of Ω, a finite number of elements $g_j \in \Omega (j = 1, 2, \cdots, n)$, their corresponding points $z_j \in (AM)$ $(j = 1, 2, \cdots, n)$ and a system of neighborhoods $V(g_j)$ $(j = 1, 2, \cdots, n)$ such that $\Omega \subseteq \sum_{j=1}^{n} V(g_j), f(z_j) > X_0(f) - \epsilon$ for any $f \in \Omega$ and $f(z_j) < X_0(f) + \epsilon$ for any $f \in V(g_j)$. Consequently, if we put $x_\epsilon = z_1 \wedge z_2 \wedge \cdots \wedge z_n$, then we have $X_0(f) - \epsilon < f(x_\epsilon) = \min_{1 \leq j \leq n} f(z_j) < X_0(f) + \epsilon$ for any $f \in \Omega$.

Thus we have proved that there exists, for any $\epsilon > 0$, a point $x_\epsilon \in (AM)$ such that $| f(x_\epsilon) - X_0(f) | < \epsilon$ for any $f \in \Omega$. Let us now consider a sequence $\{x_{\epsilon_n}\}$ $(n = 1, 2, \cdots)$, where $\epsilon_n > 0$, $\epsilon_n \to 0$. Then we have $| f(x_{\epsilon_m} - x_{\epsilon_n}) | < | f(x_{\epsilon_m}) - X_0(f) | + | f(x_{\epsilon_n}) - X_0(f) | < \epsilon_m + \epsilon_n$ for any $f \in \Omega$ and for any m and n. Consequently, by Lemma 7.1, we must have $\| x_{\epsilon_m} - x_{\epsilon_n} \| < \epsilon_m + \epsilon_n$ for any m and n. Thus $\{x_{\epsilon_n}\}$ $(n = 1, 2, \cdots)$ is a fundamental sequence in (AM), and it is clear that $x_0 = \lim_{n \to \infty} x_{\epsilon_n}$ will satisfy $f(x_0) = X_0(f)$ for any $f \in \Omega$. The proof of Lemma 7.1 and consequently the proof of Theorem 1 is completed.

REMARK. It will be easily seen that we can also state Lemma 7.2 in the following form:

THEOREM 3. *Let Ω be a compact Hausdorff space, and let E be a linear subspace of the space $C(\Omega)$ of all bounded continuous real-valued functions defined on Ω, which is closed in $C(\Omega)$ both in the sense of topology and in the sense of lattice. (E is closed in $C(\Omega)$ in the sense of lattice if $x(t), y(t) \in E$ imply* $\max (x(t), y(t))$, $\min (x(t), y(t)) \in E$.) *If we now denote by $\{t_\alpha, t'_\alpha ; \lambda_\alpha ; \alpha \in \mathfrak{M}\}$ (where $t_\alpha \in \Omega$, $t'_\alpha \in \Omega$, $0 \leqq \lambda_\alpha \leqq 1$, $(\alpha \in \mathfrak{M})$, and $\mathfrak{M}$ is a set of indices α whose power can be arbitrarily large) the system of all pairs t_α, t'_α, λ_α such that*

$$x(t_\alpha) = \lambda_\alpha x(t'_\alpha), \tag{10.2}$$

for all $x(t) \in E$, then conversely, every function $x(t) \in C(\Omega)$ which satisfies (10.2) *for any $\alpha \in \mathfrak{M}$ belongs to E; i.e., E is the subspace $C(\Omega; t_\alpha, t'_\alpha ; \lambda_\alpha ; \alpha \in \mathfrak{M})$.*

11. Proof of Theorem 2

LEMMA 11.1. *For every positive bounded linear functional $f(x)$ defined on an abstract (M)-space with a unit element $\mathbf{1}$, we have $\| f \| = f(\mathbf{1})$.*

PROOF. *Let $\epsilon > 0$* be an arbitrary positive number. Then there exists an $x \in (AM)$ such that $\| x \| = 1$ and $f(x) > \| f \| - \epsilon$. Since $\| x \| = 1$ implies $x \leqq \mathbf{1}$ by definition, we have, by the positiveness of $f(x)$, $\| f \| - \epsilon < f(x) \leqq f(\mathbf{1}) \leqq \| f \| \cdot \| \mathbf{1} \| = \| f \|$. Since $\epsilon > 0$ is arbitrary, we must have $\| f \| = f(\mathbf{1})$.

Now we shall proceed to the proof of Theorem 2, which is quite easy. We have only to show that, if there exists a unit element $\mathbf{1}$, in the given abstract (M)-space (AM), then there exist no relations of the form (7.8). Indeed, we can prove more precisely that, under the same assumption, we have $\Omega = \Omega'$ i.e., Ω' itself is compact with respect to the weak topology of $(AM)^*$.

By Lemma 11.1, every bounded linear functional $f(x) \in \Omega'$ satisfies $f(\mathbf{1}) = 1$. Hence, by Lemma 6.3, every bounded linear functional $f(x) \in \Omega = \overline{\Omega'}$ must also satisfy $f(\mathbf{1}) = 1$ and this in turn implies (again by Lemma 11.1) $\| f \| = 1$ for

any $f(x) \in \Omega$. Consequently, by the definition of Ω, we must have $\Omega = \Omega'$. The proof of Theorem 2 is completed.

We can also state the following

THEOREM 4. *Let Ω be a compact Hausdorff space, and let E be a linear subspace of the space $C(\Omega)$ of all bounded continuous real-valued functions defined on Ω which is closed in $C(\Omega)$ both in the sense of topology and lattice (see Theorem 3). If, moreover, E contains a constant function, and if there exists for any t_1, $t_2 \in \Omega$, $t_1 \neq t_2$, a continuous function $x(t) \in E$ such that $x(t_1) \neq x(t_2)$, then E must be identical with the entire space $C(\Omega)$.*

III. APPLICATIONS

12. Theorem of E. Čech

A Hausdorff space Ω is called *completely regular* if there exists, for any point $t_0 \in \Omega$ and for any neighborhood $U(t_0)$ of t_0, a continuous real-valued function $x(t)$ defined on Ω such that $x(t_0) = 1$ and $x(t) = 0$ for any $t \bar{\in} U(t_0)$.

THEOREM 5. *For any completely regular Hausdorff space Ω', there exists a compact Hausdorff space Ω which contains Ω' as a dense subset and such that every bounded continuous real-valued function defined on Ω' can be uniquely extended to a continuous function defined on Ω.*

This theorem is due to E. Čech [4].

PROOF. Consider the space $C(\Omega')$ of all bounded continuous real-valued functions $x(t)$ defined on Ω'. $C(\Omega')$ is an abstract (M)-space with respect to the norm: $\| x \| = \text{l.u.b.}_{t \in \Omega'} | x(t) |$ and the partial ordering $x \geqq y$: if and only if $x(t) \geqq y(t)$ for all $t \in \Omega'$. Moreover, $x(t) \equiv 1$ is a unit element of $C(\Omega')$. Consequently, by Theorem 2, there exists a compact Hausdorff space Ω such that $C(\Omega')$ is isometric and lattice isomorphic to the space $C(\Omega)$ of all bounded continuous real-valued functions $x(f)$ defined on Ω. (We denote elements of Ω' and Ω by t and f respectively.)

We shall show that Ω' can be homeomorphically embedded into Ω. To prove this, let t_0 be an arbitrary point of Ω', and put $f_0(x) = x(t_0)$ for all $x(t) \in C(\Omega')$. Then $f_0(x)$ is a bounded linear functional defined on $C(\Omega')$ and clearly satisfies the conditions (7.1)–(7.7) stated in section 7. Hence $f_0(x)$ can be considered as a point of Ω. Thus the mapping $t_0 \to \varphi(t_0) = f_0$ of Ω' onto $\varphi(\Omega') \subseteq \Omega$ is defined, and all that we have to do is to prove that $\varphi(\Omega')$ is dense in Ω and that the mapping $\Omega' \leftrightarrow \varphi(\Omega')$ is a homeomorphism.

In the first place, it is clear that this mapping $t \to \varphi(t) = f$ is continuous. Indeed, let $U(f_0 ; x_1, x_2, \cdots, x_n ; \epsilon)$ be an arbitrary neighborhood of a point $f_0 = \varphi(t_0) \in \varphi(\Omega')$, where $x_i \in C(\Omega')$ for $i = 1, 2, \cdots, n$, and $\epsilon > 0$. (See the definition of weak topology given in section 6.) Since every $x_i(t)$ is continuous on Ω', there exists a neighborhood $U_i(t_0)$ of t_0 such that $t \in U_i(t_0)$ implies $| x_i(t) - x_i(t_0) | < \epsilon$. Consequently, if we put $U(t_0) = \prod_{i=1}^{n} U_i(t_0)$, then $t \in U(t_0)$ implies $| x_i(t) - x_i(t_0) | < \epsilon$ for $i = 1, 2, \cdots, n$, or equivalently $f = \varphi(t) \in U(f_0 ; x_1, x_2, \cdots, x_n ; \epsilon)$, which means that the mapping $t \to \varphi(t) = f$ is continuous at $t = t_0$.

Next, it is clear that the mapping $t \to \varphi(t) = f$ is one-to-one. For, by the complete regularity of Ω', there exists, for any two different points t_0, $t_0' \in \Omega'$, $t_0 \neq t_0'$, a bounded continuous real-valued function $x_0(t)$ such that $x_0(t_0) \neq x_0(t_0')$. To prove that the inverse mapping $f \to t = \varphi^{-1}(f)$ is also continuous, let $U(t_0)$ be an arbitrary neighborhood of a point $t_0 \in \Omega'$. Then, again by the complete regularity of Ω', there exists a bounded continuous real-valued function $x_0(t)$ such that $x_0(t_0) = 1$ and $x_0(t) = 0$ for any $t \bar{\in} U(t_0)$. Consequently, every $t \in \Omega'$ which satisfies $\varphi(t) \in U(f_0 ; x_0 ; 1)$, where $f_0 = \varphi(t_0)$, must necessarily be inside of $U(t_0)$ (see the definition of weak topology given in section 6), and this fact means that $f \to t = \varphi^{-1}(f)$ is continuous at $f = f_0$.

Thus we have proved that $\Omega' \leftrightarrow \varphi(\Omega') \subseteq \Omega$ is a homeomorphic mapping. It is easy to see that the adjoint mapping $x \leftrightarrow \bar{\varphi}(x) = \bar{x}$; $\bar{x}(f) = x(\varphi(t))$ is an isometric and lattice-isomorphic mapping between $C(\Omega')$ and $C(\Omega)$. In order to prove that $\varphi(\Omega')$ is dense in Ω, let us assume that the closure $\overline{\varphi(\Omega')}$ of $\varphi(\Omega')$ is a proper subset of Ω. Then there exists a continuous function $x_0(f)$ which is zero at every point of $\overline{\varphi(\Omega')}$ without vanishing identically on Ω. This is, however, a contradiction, since on the one hand $x_0(f)$ must correspond to a function $x_0(t)$ which is identically zero on Ω' (by the mapping $\bar{\varphi}$), while on the other hand $x_0(f)$ must correspond to a function which is not identically zero on Ω' (by the isometric and lattice-isomorphic mapping of $C(\Omega')$ and $C(\Omega)$).

Thus we have seen that $\varphi(\Omega')$ is dense in Ω. Since the possibility of extension and the uniqueness of extension are both clear from our construction, the proof of Theorem 5 is completed.

REMARK 1. It is to be noted that in the proof given above we have made no essential use of Theorem 2. Indeed, we have only to consider each point $t_0 \in \Omega'$ as a bounded linear functional $f_0(x) = x(t_0)$ defined on $C(\Omega')$. The mapping $t_0 \to \varphi(t_0) = f_0$ thus obtained is a homeomorphic embedding of Ω' into the conjugate space $C^*(\Omega')$ of $C(\Omega')$. This can be proved in exactly the same way as above, and the closure $\Omega = \overline{\varphi(\Omega')}$ of the image of $\varphi(\Omega')$ of Ω' with respect to the weak topology of $C^*(\Omega')$ is the required compact set. It is easy to see that the extension of bounded continuous functions is possible and that this extension is unique.[10]

[10] Let G be a topological group and let $B(G)$ be the space of all real valued bounded continuous almost periodic functions $x(t)$ defined on G. $B(G)$ is a Banach space with respect to the norm $\| x \| = \sup_{t \in G} | x(t) |$. It is easy to see that $B(G)$ is an abstract (M)-space with a unit with respect to the ordinary partial ordering $x \geqq y$: $x(t) \geqq y(t)$ for all $t \in G$. (In this case $z = x \vee y$ is exactly the max. $(x(t), y(t))$ at each point of G). Hence by Theorem 2, there exists a compact Hausdorff space G^* such that $B(G)$ is isometric and lattice isomorphic to the space $C(G^*)$ of all real valued continuous functions $x(t)$ defined on G^*. It is not difficult to see that in this case G^* is even a topological group, and that, in case G has sufficiently many almost periodic functions, G is a dense subgroup of G^*. Thus, for every topological group G with sufficiently many almost periodic functions, there exists a compact topological group G^* which contains G as a dense subgroup and such that every continuous almost periodic function defined on G can be uniquely extended to a continuous (and hence almost periodic) function defined on G^*. This is also a conse-

REMARK 2. There are two extreme cases of completely regular Hausdorff spaces, which seem to be interesting, namely, the case of compact Hausdorff spaces and the case of discrete spaces. These will be discussed in the following sections.

13. Case of compact Hausdorff spaces

In the arguments given above in section 12, consider the special case when the given Hausdorff space Ω' is compact. Then the homeomorphic image $\varphi(\Omega')$ of Ω' must also be compact, and consequently Ω must coincide with $\varphi(\Omega')$; i.e. Ω is homeomorphic to Ω'. This means that, if we start from a compact Hausdorff space Ω' and consider the abstract (M)-space $C(\Omega')$ of all bounded continuous real-valued functions defined on Ω', then the compact Hausdorff space Ω obtained in Theorem 2 is homeomorphic to the original space Ω', or in other words:

THEOREM 6. *The compact Hausdorff space Ω obtained in Theorem 2 is uniquely determined up to a homeomorphism.*

We can also state this theorem in the following form:

THEOREM 7. *Let Ω_1 and Ω_2 be two compact Hausdorff spaces. Then, in order that the two abstract (M)-spaces $C(\Omega_1)$ and $C(\Omega_2)$ be isometric and lattice-isomorphic, it is necessary and sufficient that Ω_1 and Ω_2 be homeomorphic to each other.*

Incidentally, we have also proved the following

THEOREM 8. *Let Ω be a compact Hausdorff space and let $C(\Omega)$ be a Banach space of all bounded continuous real-valued functions $x(t)$ defined on Ω. Then, every bounded linear functional $f_0(x)$ defined on $C(\Omega)$ which satisfies the conditions:*

$$\|f_0\| = 1, \tag{13.1}$$

$$x \geqq 0 \quad \textit{implies} \quad f_0(x) \geqq 0, \tag{13.2}$$

$$x \wedge y = 0 \quad \textit{implies} \quad f_0(x) = 0 \quad \textit{or} \quad f_0(y) = 0, \tag{13.3}$$

(i.e., *all conditions* (7.1)–(7.7)), *can be uniquely expressed in the form*:

$$f_0(x) = x(t_0), \tag{13.4}$$

where t_0 is a point of Ω.

Let us now consider a completely additive non-negative set function $\mu(E)$ defined for all Borel sets E of Ω such that $\mu(\Omega) = 1$. Then the integral

$$f_\mu(x) = \int_\Omega x(t)\mu(dt) \tag{13.5}$$

---

quence of a general duality theorem of Tannaka (Dualität der nichtkommutativen Gruppen, Tohoku Math. Journ., 45(1939), 1–12). The compactification of G with respect to a single almost periodic function (in this case we have to take a factor group G/N of G by a certain invariant subgroup N of G and then make this factor group G/N into a compact group $(G/N)^*$) was discussed by A. Weil (Sur les fonctions presque périodiques de von Neumann, C. R. Paris, 200(1935), 38–40) and the general case may be obtained from this by applying the idea of uniform space. (As for uniform spaces see A. Weil, Sur les espaces a structure uniforme et sur la topologie générale, Actualité, 551, Paris, 1937). It may be noticed that, after all, the method of uniform space and that of weak topology amount to the same thing.

will define a bounded linear functional $f_\mu(x)$ defined on $C(\Omega)$. It is easy to see that this $f_\mu(x)$ satisfies the conditions (13.1) and (13.2).

Moreover, we can prove that conversely every such functional is of the form (13.5); i.e.,

THEOREM 9. *Let Ω be a compact Hausdorff space, and let $C(\Omega)$ be defined as usual. Then every bounded linear functional $f(x)$ defined on $C(\Omega)$ which satisfies the conditions* (13.1) *and* (13.2) *can be expressed in the form* (13.5), *where $\mu(E)$ is a completely additive non-negative set funtion defined for all Borel sets E of Ω such that $\mu(\Omega) = 1$.*

This theorem was proved by S. Saks [15] under the assumption that Ω is a compact metric space.

PROOF.[11] For any open set O of Ω, put

$$\mu(O) = \text{l.u.b. } f(x), \tag{13.6}$$

where l.u.b. means the least upper bound for all continuous functions $x(t) \in C(\Omega)$ such that $x(t) \leqq 1$ for $t \in O$ and $x(t) = 0$ for $t \in \Omega - O$. It is clear that $\mu(\Omega) = 1$ and that $O \subseteq P$ implies $\mu(O) \leqq \mu(P)$ for any open sets O and P. It is also not difficult to see that we have $\mu(O + P) \leqq \mu(O) + \mu(P)$[12] for any open sets O and P, where the equality holds if O and P are disjoint. Next, for any subset M of Ω, put

$$\mu^*(M) = \text{g.l.b. } \mu(O), \tag{13.7}$$

where g.l.b. means the greatest lower bound for all open sets O which contain M. It is clear that for open sets O we have $\mu^*(O) = \mu(O)$. We shall first prove that $\mu^*(M)$ is a Carathéodory outer measure and that every open set is μ^*-measurable (in the sense of Carathéodory). Among the axioms of Carathéodory we have only to prove that

$$\mu^*(\textstyle\sum_{n=1}^{\infty} M_n) \leqq \sum_{n=1}^{\infty} \mu^*(M_n) \tag{13.8}$$

for any sequence of subsets $\{M_n\}$ $(n = 1, 2, \cdots)$ of Ω such that $\sum_{n=1}^{\infty} \mu^*(M_n) < \infty$. It is easy to see that we have only to prove (13.8) for the special case when all M_n are open, i.e.,

$$\mu(\textstyle\sum_{n=1}^{\infty} O_n) \leqq \sum_{n=1}^{\infty} \mu(O_n) \tag{13.9}$$

---

[11] The proof given below follows the idea of Professor von Neumann. See his lecture note on Haar measure [14].

[12] Let $\epsilon > 0$ be an arbitrary positive number. Then there exists an $x(t) \in C(\Omega)$ such that $f(x) > \mu(O + P) - \epsilon$, $x(t) \leqq 1$ for $t \in O + P$ and $x(t) = 0$ for $t \in \Omega - (O + P)$. Let C be the set of all $t \in \Omega$ at which $x(t) \geqq \epsilon$. Then C is a closed set contained in $O + P$. Hence C is a sum of two closed (not necessarily disjoint) sets C_1 and C_2 : $C = C_1 + C_2$ such that $C_1 \subseteq O$ and $C_2 \subseteq P$. Consequently, there exist two continuous functions $y(t)$, $z(t) \in C(\Omega)$ such that $y(t) = 1$ for $t \in C_1$, $0 \leqq y(t) \leqq 1$ for $t \in O - C_1$, $y(t) = 0$ for $t \in \Omega - O$; $z(t) = 1$ for $t \in C_2$, $0 \leqq z(t) \leqq 1$ for $t \in P - C_2$, $z(t) = 0$ for $t \in \Omega - P$. Then it is clear that $x(t) \leqq y(t) + z(t) + \epsilon$ for all $t \in \Omega$ and $f(y) \leqq \mu(O)$, $f(z) \leqq \mu(P)$. Consequently, $\mu(O + P) - \epsilon < f(x) \leqq f(y) + f(z) + \epsilon \leqq \mu(O) + \mu(P) + \epsilon$. Since $\epsilon > 0$ is arbitrary this proves the required inequality $\mu(O + P) \leqq \mu(O) + \mu(P)$.

for any sequence of open sets $\{O_n\}$ $(n = 1, 2, \cdots)$ such that $\sum_{n=1}^{\infty}\mu(O_n) < \infty$. Indeed, for any $\epsilon > 0$, let O_n be an open set such that $O_n \supseteq M_n$ and $\mu(O_n) < \mu^*(M_n) + 2^{-n}\cdot\epsilon$. Then we have $\sum_{n=1}^{\infty}O_n \supseteq \sum_{n=1}^{\infty}M_n$ and $\mu^*(\sum_{n=1}^{\infty}M_n) \leqq \mu(\sum_{n=1}^{\infty}O_n) \leqq \sum_{n=1}^{\infty}\mu(O_n) < \sum_{n=1}^{\infty}\mu^*(M_n) + \epsilon$. Since $\epsilon > 0$ is arbitrary we have (13.8).

Now, in order to prove (13.9), let $\epsilon > 0$ be an arbitrary positive number. Then there exists an $x(t) \in C(\Omega)$ such that $f(x) > \mu(\sum_{n=1}^{\infty}O_n) - \epsilon$, $x(t) \leqq 1$ for $t \in \sum_{n=1}^{\infty}O_n$ and $x(t) = 0$ for $t \in \Omega - \sum_{n=1}^{\infty}O_n$. If we denote by C the set of all points t at which $x(t) \geqq \epsilon$, then C is a closed set contained in $\sum_{n=1}^{\infty}O_n$. Hence there exists an integer N such that $C \subseteq \sum_{n=1}^{N}O_n$. Let now $y(t) \in C(\Omega)$ be such that $y(t) = 1$ for $t \in C$, $y(t) = 0$ for $t \in \Omega - \sum_{n=1}^{N}O_n$, and $0 \leqq y(t) \leqq 1$ elsewhere in Ω. Then we have $f(y) \leqq \mu(\sum_{n=1}^{N}O_n)$ and $x(t) \leqq y(t) + \epsilon$ for all $t \in \Omega$. Hence we have $\mu(\sum_{n=1}^{\infty}O_n) - \epsilon < f(x) \leqq f(y) + \epsilon \leqq \mu(\sum_{n=1}^{N}O_n) + \epsilon \leqq \sum_{n=1}^{N}\mu(O_n) + \epsilon \leqq \sum_{n=1}^{\infty}\mu(O_n) + \epsilon$. Since $\epsilon > 0$ is arbitrary we have (13.9).

In order to prove that every open set is μ^*-measurable, we have only to show that

$$\mu^*(M) \geqq \mu^*(M\cdot O) + \mu^*(M - M\cdot O) \tag{13.10}$$

for any open set O and for any subset $M \subseteq \Omega$. It is again easy to see that we have only to prove that (13.10) is true when M is open, i.e.

$$\mu(P) \geqq \mu(P\cdot O) + \mu^*(P - P\cdot O) \tag{13.11}$$

for any open sets O and P. Indeed, for any $\epsilon > 0$, there exists an open set $P \supseteq M$ such that $\mu(P) < \mu^*(M) + \epsilon$. Then we have $\mu^*(M) + \epsilon > \mu(P) \geqq \mu(P\cdot O) + \mu^*(P - P\cdot O) \geqq \mu^*(M\cdot O) + \mu^*(M - M\cdot O)$, and since $\epsilon > 0$ is arbitrary, we have (13.10).

To prove (13.11), let $\epsilon > 0$ be an arbitrary positive number. Then there exists an $x(t) \in C(\Omega)$ such that $f(x) > \mu(P\cdot O) - \epsilon$, $x(t) \leqq 1$ for $t \in P\cdot O$ and $x(t) = 0$ for $t \in \Omega - P\cdot O$. Let C be a set of all points t at which $x(t) \geqq \epsilon$. Then C is a closed set contained in $P\cdot O$. Let now $y(t) \in C(\Omega)$ be such that $f(y) > \mu(P - C) - \epsilon$, $y(t) \leqq 1$ for $t \in P - C$ and $y(t) = 0$ for $t \in \Omega - (P - C) = C + (\Omega - P)$. Then we have $x(t) + y(t) \leqq 1 + \epsilon$ for $t \in P$ and $x(t) + y(t) = 0$ for $t \in \Omega - P$. Hence we have $f(x) + f(y) \leqq \mu(P) + \epsilon$ and consequently $\mu^*(P - P\cdot O) + \mu(P\cdot O) - 2\epsilon \leqq \mu(P - C) + \mu(P\cdot O) - 2\epsilon < f(x) + f(y) < \mu(P) + \epsilon$. Since $\epsilon > 0$ is arbitrary we have (13.11).

Thus we have proved that μ^* is a Carathéodory outer measure and every open set is μ^*-measurable. Consequently every Borel set is μ^*-measurable, and if we put $\mu(E) = \mu^*(E)$ for Borel sets E, then $\mu(E)$ is the required measure. Since it is clear that $\mu(E)$ is completely additive and that $\mu(\Omega) = 1$, we have only to prove that we have

$$f(x) = \int_{\Omega} x(t)\mu(dt) \tag{13.12}$$

for any continuous function $x(t) \in C(\Omega)$.

Let $\epsilon > 0$ be an arbitrary positive number and let $\alpha_0 < \alpha_1 < \alpha_2 < \cdots < \alpha_n$ be a system of real numbers such that $\alpha_0 < -\| x \|$, $\alpha_n > \| x \|$ and $\alpha_i - \alpha_{i-1} < \epsilon$ for $i = 1, 2, \cdots, n$. Let E_i be the set of points t at which $x(t) > \alpha_i (i = 1, 2, \cdots, n)$. Then, each E_i is open, and if we put

$$x_i(t) \begin{cases} = 0, & \text{if } x(t) < \alpha_{i-1}, \\ = (\alpha_i - \alpha_{i-1})^{-1}(x(t) - \alpha_{i-1}), & \text{if } \alpha_{i-1} \leqq x(t) \leqq \alpha_i, \\ = 1 & \text{if } x(t) > \alpha_i, \end{cases}$$

then $x_i(t) \leqq 1$ for $t \in E_{i-1}$, $x_i(t) = 0$ for $t \in \Omega - E_{i-1}$ and $x(t) = \alpha_0 + \sum_{i=1}^{n} (\alpha_i - \alpha_{i-1})x_i(t)$. Hence we have (by (13.6)) $f(x_i) \leqq \mu(E_{i-1})$ and consequently (since $\mu(E_0) = 1$ and $\mu(E_n) = 0$)

$$\begin{aligned} f(x) &= \alpha_0 + \textstyle\sum_{i=1}^{n}(\alpha_i - \alpha_{i-1})f(x_i) \\ &\leqq \alpha_0 + \textstyle\sum_{i=1}^{n}(\alpha_i - \alpha_{i-1})\mu(E_{i-1}) \\ &= \textstyle\sum_{i=1}^{n}\alpha_i(\mu(E_{i-1}) - \mu(E_i)) \\ &\leqq \textstyle\sum_{i=1}^{n}\alpha_{i-1}(\mu(E_{i-1}) - \mu(E_i)) + \epsilon \\ &\leqq \int_{\Omega} x(t)\mu(dt) + \epsilon. \end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we have $f(x) \leqq \int_{\Omega} x(t)\mu(dt)$. Since the inverse inequality follows by replacing $x(t)$ by $-x(t)$, the proof of Theorem 9 is completed.

REMARK 1. J. v. Neumann [14] started from a set function $\lambda(C)$ defined for compact sets C, and defined $\mu(O)$ for open sets O as follows:

$$\mu(O) = \text{l.u.b. } \lambda(C), \tag{13.13}$$

where l.u.b. means the least upper bound for all compact sets C contained in O. In our case, we can, indeed, define $\lambda(C)$ by

$$\lambda(C) = \text{g.l.b. } f(x), \tag{13.14}$$

where g.l.b. means the greatest lower bound for all $x(t) \in C(\Omega)$ such that $x(t) = 1$ for $t \in C$ and $x(t) \geqq 0$ for $t \in \Omega - C$. Then it is easy to see that the formula (13.13) gives the same value of $\mu(O)$ as (13.6).

REMARK 2. A completely additive non-negative set-function $\mu(E)$ is called *regular*, if there exist, for any Borel set E and for any $\epsilon > 0$ a closed set $C \subseteq E$ and an open set $O \supseteq E$ such that $\mu(C) > \mu(E) - \epsilon$, $\mu(O) < \mu(E) + \epsilon$. (If $\mu(E)$ is not non-negative, this last condition must be replaced by $|\mu(E) - \mu(E')| < \epsilon$ for any Borel set E' such that $C \subseteq E' \subseteq O$.)

In Theorem 9, the uniqueness of $\mu(E)$ was not discussed. If, however, we require that $\mu(E)$ be regular, then $\mu(E)$ is uniquely determined. Indeed, if there exist two completely additive regular set-functions $\mu_1(E)$ and $\mu_2(E)$ which satisfy (13.12) for any $x(t) \in C(\Omega)$, then $\mu(E) = \mu_1(E) - \mu_2(E)$ is also a completely additive regular set-function which satisfies

$$\int_{\Omega} x(t)\mu(dt) = 0$$

for any continuous function $x(t) \in C(\Omega)$. We shall show that we have $\mu(E) = 0$ for any Borel set E. Let E be an arbitrary Borel set. Then, for any $\epsilon > 0$, there exist a closed set $C \subseteq E$ and an open set $O \supseteq E$ such that $|\mu(E) - \mu(E')| < \epsilon$ for any Borel set E' with $C \subseteq E' \subseteq O$. Let now $x(t) \in C(\Omega)$ be a continuous function such that $x(t) = 1$ for $t \in C$, $x(t) = 0$ for $t \in \Omega - O$, and $0 \leqq x(t) \leqq 1$ elsewhere in Ω. Then, denoting by $x_E(t)$ the characteristic function of E, we have

$$\begin{aligned} |\mu(E)| &= \left|\mu(E) - \int_\Omega x(t)\mu(dt)\right| \\ &= \left|\int_\Omega (x_E(t) - x(t))\mu(dt)\right| \\ &\leqq \int_{O-C} |x_E(t) - x(t)|\,|\mu(dt)| \\ &\leqq \int_{O-C} |\mu(dt)| \leqq \epsilon. \end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we must have $\mu(E) = 0$ for any Borel set E, and the uniqueness of $\mu(E)$ in Theorem 9 is proved.

Since, as is easily seen, every bounded linear functional defined on $C(\Omega)$ can be expressed as a difference of two positive bounded linear functionals, the result obtained in Theorem 9 can be extended to a following form:

THEOREM 10. *Let Ω be a compact Hausdorff space, and let $C(\Omega)$ be defined as usual. Let us further denote by $\mathfrak{M}(\Omega)$ the space of all completely additive regular set-functions $\mu(E)$ defined for all Borel sets E of Ω. If we put $\|\mu\|$ = total variation of $\mu(E)$ = l.u.b.$_{E\subseteq\Omega}\,\mu(E)$ − g.l.b.$_{E\subseteq\Omega}\,\mu(E)$, and $\mu \geqq \nu$ if and only if $\mu(E) \geqq \nu(E)$ for any Borel set $E \subseteq \Omega$, then $\mathfrak{M}(\Omega)$ is isometric and lattice isomorphic to the conjugate space $C^*(\Omega)$ of $C(\Omega)$. The correspondence between $C^*(\Omega)$ and $\mathfrak{M}(\Omega)$ is given by* (13.12).

We omit the proof.

EXAMPLE 1. Let Ω_0 be a closed interval: $0 \leqq t \leqq 1$. Then, as is well known, $C^*(\Omega_0)$ is a space of all functions of bounded variation (with total variation as its norm). It is easy to see that this space is nothing but the space $\mathfrak{M}(\Omega_0)$ defined in Theorem 10.

EXAMPLE 2. Let Ω be a compact Hausdorff space consisting of $t_1, t_2, \cdots, t_n, \cdots$ and t_∞, where each t_n $(n = 1, 2, \cdots)$ is isolated and t_∞ is a limiting point of all others. Then $C(\Omega)$ is the space (C_0) of all convergent sequences of real numbers $x = (x_1, x_2, \cdots, x_n, \cdots, \lim_n x_n = x_\infty)$, with $\|x\| = \text{l.u.b.}_n |x_n|$ as its norm. In this case $C^*(\Omega)$ is the space of all sequences of real numbers $f = (f_1, f_2, \cdots, f_n, \cdots, f_\infty)$ such that $\|f\| = \sum_{n=1}^{\infty} |f_n| + |f_\infty| < \infty$.

In Theorem 9 we have seen that every positive bounded linear functional defined on $C(\Omega)$ can be considered as a measure $\mu(E)$ defined for all Borel sets E of Ω. And, as is easily seen, Theorem 8 means that if we assume the additional

condition (13.3) then this measure $\mu(E)$ is concentrated in one point t_0 of Ω. Moreover, we can show the following:

THEOREM 11. *Let Ω be a compact Hausdorff space and let $C(\Omega)$ be a space of all bounded continuous real-valued functions $x(t)$ defined on Ω. Then every bounded linear functional $f_\mu(x)$ of the form* (13.5) *can be approximated (in the sense of weak topology) by convex combinations $f_\Sigma(x)$ of the functionals of the form* (13.4):

$$f_\Sigma(x) = \sum_{p=1}^{k} \alpha_p x(t_p), \tag{13.15}$$

where $t_p \in \Omega$, $\alpha_p \geqq 0$ $(p = 1, 2, \cdots, k)$ and $\sum_{p=1}^{k} \alpha_p = 1$.

In other words, if we denote by Δ_0 and Δ_μ the set of all bounded linear functionals $f_0(x)$ and $f_\mu(x)$ of the form (13.4) *and* (13.5) *respectively, then Δ_μ is contained in (hence coincides with) the convex closure of Δ_0 with respect to the weak topology.*

PROOF. Let $f_\mu(x)$ be a bounded linear functional which is defined by (13.5), where $\mu(E)$ is a completely additive non-negative set-function defined for all Borel sets E of Ω with $\mu(\Omega) = 1$. Then, for any $x(t) \in C(\Omega)$ and for any $\epsilon > 0$, there exists a decomposition $\Omega = \sum_{p=1}^{k} E_p$, $E_p \cdot E_q = 0$ $(p \neq q)$ of Ω into disjoint Borel sets such that

$$|f_\mu(x) - \sum_{p=1}^{k} x(t_p)\mu(E_p)| < \epsilon$$

for any $t_p \in E_p$, $p = 1, 2, \cdots, k$. Let now $x_i(t) \in C(\Omega)$ $(i = 1, 2, \cdots, n)$ be an arbitrary finite system and let $\Omega = \sum_{p=1}^{k_i} E_p^{(i)}$, $E_p^{(i)} \cdot E_q^{(i)} = 0$ $(p \neq q)$ be the corresponding decomposition of Ω into Borel sets (with the same ϵ), then the decomposition: $\Omega = \sum_{p_1=1}^{k_1} \sum_{p_2=1}^{k_2} \cdots \sum_{p_n=1}^{k_n} E_{p_1 p_2 \cdots p_n}$, where $E_{p_1 p_2 \cdots p_n} = E_{p_1}^{(1)} E_{p_2}^{(2)} \cdots E_{p_n}^{(n)}$, will satisfy

$$|f_\mu(x_i) - \sum_{p_1=1}^{k_1} \sum_{p_2=1}^{k_2} \cdots \sum_{p_n=1}^{k_n} x(t_{p_1 p_2 \cdots p_n})\mu(E_{p_1 p_2 \cdots p_n})| < \epsilon$$

for any $t_{p_1 p_2 \cdots p_n} \in E_{p_1 p_2 \cdots p_n}$ $(1 \leqq p_i \leqq k_i, i = 1, 2, \cdots, n)$ and for any x_i $(i = 1, 2, \cdots, n)$. If we now put $\alpha_{p_1 p_2 \cdots p_n} = \mu(E_{p_1 p_2 \cdots p_n})$, then this fact means that the bounded linear functional

$$f_\Sigma(x) = \sum_{p_1=1}^{k_1} \sum_{p_2=1}^{k_2} \cdots \sum_{p_n=1}^{k_n} \alpha_{p_1 p_2 \cdots p_n} f(t_{p_1 p_2 \cdots p_n}) \tag{13.16}$$

lies inside a weak neighborhood $U(f_\mu ; x_1, x_2, \cdots, x_n ; \epsilon)$ of f_μ. Consequently, f_μ is a limiting point of all functionals $f_\Sigma(x)$ of the form (13.7). Since $1 = \sum_{p_1=1}^{k_1} \sum_{p_2=1}^{k_2} \cdots \sum_{p_n=1}^{k_n} \alpha_{p_1 p_2 \cdots p_n}$, this proves Theorem 11.

REMARK. This theorem is also true even if Ω is not compact.

14. Case of discrete spaces, 1. Theorem of M. H. Stone

In this section we shall consider the case when Ω' is discrete, i.e., when every point of Ω' is an isolated point. Any two such discrete spaces with the same power are clearly homeomorphic to each other. Consequently, the compact Hausdorff space Ω which is obtained in Theorem 5, is uniquely determined (up to a homeomorphism) by the power $\mathfrak{m}$ of the discrete space Ω. We shall denote this space by $\Omega(\mathfrak{m})$.

Since the space Ω' is discrete by assumption, every bounded real-valued function is continuous on Ω'. Let now $x_E(t)$ be a characteristic function of an arbitrary subset E of Ω'. Then $x_E(t)$ is continuous on Ω' and consequently $x_E(t)$ can be uniquely extended to a continuous function $x_E(f)$ defined on Ω. It is clear that this function $x_E(f)$ takes only the values 0 and 1. Indeed, since $\varphi(\Omega')$ is dense in Ω, there exists, for any $\epsilon > 0$, a point $t \in \Omega'$ such that $| x_E(f) - x_E(t) | < \epsilon$. Consequently, since $x_E(t)$ takes only the values 0 and 1, we must have $x_E(f) = 0$ or $= 1$ for any $f \in \Omega$.

Thus we have seen that $x_E(f)$ is also a characteristic function of a certain subset of Ω. Denote this set by $\bar{E}$. It is clear that $\bar{E}$ is open-and-closed, and if we consider Ω' as a subset of Ω (this is possible since Ω' is homeomorphic to $\varphi(\Omega') \subseteq \Omega$), then we have $E \subseteq \bar{E}$. Moreover, it is easy to see that in case E is reduced to one point, $\bar{E}$ must also be reduced to the same point.

Conversely, let $\bar{E}$ be an arbitrary open-and-closed subset of Ω. Then the characteristic function $\bar{x}_{\bar{E}}(f)$ of $\bar{E}$ is continuous on Ω, and the part $x_{\bar{E}}(t)$ of $\bar{x}_{\bar{E}}(f)$ on Ω' must take only the values 0 and 1; i.e., $x_{\bar{E}}(t)$ is also a characteristic function of a certain subset of Ω'. Denote this set by E. Then it is easy to see that the correspondence $E \leftrightarrow \bar{E}$ is one-to-one. Moreover, as is easily seen, this mapping $E \leftrightarrow \bar{E}$ is isomorphic; i.e., $E \leftrightarrow \bar{E}$, $F \leftrightarrow \bar{F}$ imply $E + F \leftrightarrow \bar{E} + \bar{F}$, $E \cdot F \leftrightarrow \bar{E} \cdot \bar{F}$ and $\Omega' - E \leftrightarrow \Omega - \bar{E}$.

This result is nothing but the theorem of M. H. Stone [17] on the concrete representation of a Boolean algebra $\mathfrak{B}$ for the special case when $\mathfrak{B}$ is the Boolean algebra of all subsets of a space Ω'. But we can proceed further, as Stone did, to the general case of an abstract Boolean algebra with a unit element. (The case of a Boolean algebra without a unit element can be treated by using Theorem 1. But, in order to avoid unnecessary confusions, we shall discuss only those cases when a unit element exists.)

In the following lines we shall give a general outline of our proof. Let $\mathfrak{B}$ be an abstract Boolean algebra with a unit element e, and let us consider a space $M(\mathfrak{B})$ of all step functions defined on $\mathfrak{B}$. Under a step function defined on $\mathfrak{B}$ we understand a system $\begin{pmatrix} a_1, & a_2, & \cdots, & a_n \\ \alpha_1, & \alpha_2, & \cdots, & \alpha_n \end{pmatrix}$ where $a_i \in \mathfrak{B}$ $(i = 1, 2, \cdots, n)$, $\sum_{i=1}^{n} a_i = e$, $a_i a_j = 0$ $(i \neq j)$ and $\alpha_i (i = 1, 2, \cdots, n)$ is a real number. Two step functions $\begin{pmatrix} a_1, & a_2, & \cdots, & a_m \\ \alpha_1, & \alpha_2, & \cdots, & \alpha_m \end{pmatrix}$ and $\begin{pmatrix} b_1, & b_2, & \cdots, & b_n \\ \beta_1, & \beta_2, & \cdots, & \beta_n \end{pmatrix}$ are considered as equal if and only if $\sum_{\alpha_i = \gamma} a_i = \sum_{\beta_j = \gamma} b_j$ for any real number γ, where $\sum_{\alpha_i = \gamma} a_i$, for example, means the sum of all a_i such that $\alpha_i = \gamma$. If we now put

$$\begin{pmatrix} a_1, & a_2, & \cdots, & a_m \\ \alpha_1, & \alpha_2, & \cdots, & \alpha_m \end{pmatrix} + \begin{pmatrix} b_1, & b_2, & \cdots, & b_n \\ \beta_1, & \beta_2, & \cdots, & \beta_n \end{pmatrix} = \begin{pmatrix} a_1 b_1, & a_1 b_2, & \cdots, & a_m b_n \\ \alpha_1 + \beta_1, & \alpha_1 + \beta_2, & \cdots, & \alpha_m + \beta_n \end{pmatrix}, \tag{14.1}$$

$$\lambda \begin{pmatrix} a_1, & a_2, & \cdots, & a_n \\ \alpha_1, & \alpha_2, & \cdots, & \alpha_n \end{pmatrix} = \begin{pmatrix} a_1, & a_2, & \cdots, & a_n \\ \lambda\alpha_1, & \lambda\alpha_2, & \cdots, & \lambda\alpha_n \end{pmatrix}, \tag{14.2}$$

$$\left\| \begin{pmatrix} a_1, & a_2, & \cdots, & a_n \\ \alpha_1, & \alpha_2, & \cdots, & \alpha_n \end{pmatrix} \right\| = \max_{1 \leqq i \leqq n} |\alpha_i|, \tag{14.3}$$

$$\begin{pmatrix} a_1, & a_2, & \cdots, & a_n \\ \alpha_1, & \alpha_2, & \cdots, & \alpha_n \end{pmatrix} \geqq 0 \quad \text{if and only if } \alpha_i \geqq 0 \text{ for } i = 1, 2, \cdots, n, \tag{14.4}$$

then $M(\mathfrak{B})$ satisfies all the axioms of abstract (M)-spaces except the completeness with respect to the norm (14.3). If we now make it complete with respect to this norm, then the space $\overline{M(\mathfrak{B})}$ thus obtained is surely an abstract (M)-space with $\begin{pmatrix} e \\ 1 \end{pmatrix}$ as its unit element. (Of course, we must define the partial ordering in $\overline{M(\mathfrak{B})}$ in a suitable manner). Moreover, if we put $x_a = \begin{pmatrix} a, & e - a \\ 1, & 0 \end{pmatrix}$, then $a \leftrightarrow x_a$ is a one-to-one mapping which embeds $\mathfrak{B}$ isomorphically into $\overline{M(\mathfrak{B})}$.

If we now apply Theorem 2 to this space $\overline{M(\mathfrak{B})}$, then we have a concrete representation of $\overline{M(\mathfrak{B})}$ as the space $C(\Omega)$ of all real valued continuous functions $x(f)$ defined on a compact Hausdorff space Ω. The unit element $\mathbf{1} = \begin{pmatrix} e \\ 1 \end{pmatrix}$ of $\overline{M(\mathfrak{B})}$ is represented by a constant function $x(f) \equiv 1$, and it is easy to see that, for any $a \in \mathfrak{B}$ the function $x_a(f) \in C(\Omega)$ which corresponds to $x_a \in \overline{M(\mathfrak{B})}$ is a characteristic function of a certain subset of Ω. This follows from the fact that $2x_a \wedge \mathbf{1} = x_a$ for any $a \in \mathfrak{B}$. Let us denote this set by E_a. E_a is clearly an open-and-closed subset of Ω and the correspondence $a \leftrightarrow E_a$ is an isomorphism. Thus $\mathfrak{B}$ is concretely represented as a Boolean algebra of open-and-closed subsets of Ω. Moreover, it is easy to see that Ω is a totally disconnected space, and that for any open-and-closed subset E of Ω, there exists an $a \in \mathfrak{B}$ such that $E_a = E$.

Thus we have proved Stone's

THEOREM 12. *For any abstract Boolean algebra $\mathfrak{B}$ with a unit element, there exists a totally disconnected compact Hausdorff space Ω such that $\mathfrak{B}$ is isomorphic to the Boolean algebra of all open-and-closed subsets of Ω.*

If we, however, wish to obtain a direct proof of this theorem, then we can proceed in the following way: let $\mathfrak{B}$ be an arbitrary Boolean algebra with unit e and consider the totality Ω^* of all real valued functions $f(x)$ defined on $\mathfrak{B}$ whose values are either 0 or 1. Ω^* is a compact Hausdorff space with respect to the weak topology.[13] Let us further consider a real valued function $f(x) \in \Omega^*$ which satisfies the following conditions:

$$f(0) = 0, \quad f(e) = 1, \tag{14.5}$$

$$f(x \vee y) = \max(f(x), f(y)), \tag{14.6}$$

$$f(x \wedge y) = \min(f(x), f(y)), \tag{14.7}$$

$$f(x + y) = f(x) + f(y), \quad \text{mod } 2, \tag{14.8}$$

$$f(xy) = f(x)f(y), \tag{14.9}$$

where $x \vee y = x + y + xy$, $x \wedge y = xy$. (These conditions are clearly not mutually independent.) These conditions mean that $x \to f(x)$ is a homomorphism of $\mathfrak{B}$ onto the Boolean algebra $\mathfrak{B}_0$ consisting of 0 and 1 ($\mathfrak{B}_0$ is the simplest non-trivial Boolean algebra). The totality Ω of all such $f(x)$ constitutes again a compact Hausdorff space with respect to the weak topology,[13] and if we put $x(f) = f(x)$ for any fixed $x \in \mathfrak{B}$, then $x(f)$ is a continuous function defined on Ω whose values are always 0 or 1. It is easy to see that $x \to x(f)$ is a homomorphic mapping. In order to show that this is an isomorphism, we have only to prove that for any $a \in \mathfrak{B}$, $b \in \mathfrak{B}$, $a \neq b$, there exists an $f_0 \in \Omega$ such that $f_0(a) \neq f_0(b)$. This fact corresponds to Lemma 7.1 in the discussions of Part II, and also to the existence of a prime ideal, which contains one of a and b but not the other, in the proof of M. H. Stone.

In order to prove the existence of such a function, let Σ be an arbitrary finite subset of $\mathfrak{B}$ which contains 0, 1, a and b. Then there exists a function $f_\Sigma(x) \in \Omega^*$ which satisfies (14.5)–(14.9) for any $x, y \in \Sigma$ and such that $f_\Sigma(a) \neq f_\Sigma(b)$. This follows from the fact that the Boolean sub-algebra $\mathfrak{B}_\Sigma$ of $\mathfrak{B}$ which is generated by Σ has also only a finite number of elements, and that for such a finite Boolean algebra $\mathfrak{B}_\Sigma$ Theorem 12 is trivially true. Let Ω_Σ be the set of all such $f_\Sigma(x)$. Then Ω_Σ is clearly a non-empty closed (and consequently compact) subset of Ω^*. Moreover, the family $\{\Omega_\Sigma\}$, where Σ runs all finite subsets of $\mathfrak{B}_1$ has a finite intersection property. This follows from the fact that $\prod_{i=1}^n \Omega_{\Sigma_i} \supseteq \Omega_{\Sigma_1+\Sigma_2+\cdots+\Sigma_n} \neq 0$ for any finite number of finite subsets Σ_i ($i = 1, 2, \cdots, n$) of $\mathfrak{B}$. Hence there exists an element f_0 which belongs to all such Ω_Σ and this element f_0 is clearly the required one.

Thus we have proved that the Boolean algebra $\mathfrak{B}$ is isomorphically represented by the system of continuous functions $x(f) = f(x)$ (whose values are 0 or 1) defined on a compact space Ω. Let now E_a be the set of all $f(x) \in \Omega$ such that $f(a) = 1$. Then E_a is clearly closed, and since $\Omega - E_a = E_{e-a}$, E_a must also be open. It is easy to see that $a \leftrightarrow E_a$ is an isomorphism, and all that we have to do is to prove that Ω is totally disconnected and that for any open-and-closed subset E of Ω there exists an element $a \in \mathfrak{B}$ such that $E = E_a$. (This part of the proof corresponds to Lemma 7.2.)

To prove these, let us first show that for any $f \in \Omega$, $g \in \Omega$, $f \neq g$, there exists an open-and-closed subset E of Ω such that $f \in E$, $g \bar{\in} E$. Since $f \neq g$, there exists an $a_0 \in \mathfrak{B}$ such that $f(a_0) \neq g(a_0)$. Without loss of generality we may assume $f(a_0) = 1$, $g(a_0) = 0$, since otherwise we may replace a_0 by $e - a_0$. Then the open-and-closed subset E_{a_0} has the required property.

Next we shall show that for any $f \in \Omega$ and for any open set $U(f)$ which contains f, there exists an $a_0 \in \mathfrak{B}$ such that $f \in E_{a_0} \subseteq U(f)$. By the above result, for any $g \bar{\in} U(f)$, there exists an $a \in \mathfrak{B}$ such that $f \in E_a$, $g \bar{\in} E_a$. Since $\Omega - E_a$ is an open set which contains g, there exists by the compactness of $\Omega - U(f)$,

[13] For any $f_0 \in \Omega^*$ its weak neighborhood $U(f_0; x_1, x_2, \cdots, x_n)$ is defined as the totality of all $f \in \Omega^*$ such that $f(x_i) = f_0(x_i)$ for $i = 1, 2, \cdots, n$, where $\{x_i\}$ ($i = 1, 2, \cdots, n$) is an arbitrary finite system of elements of $\mathfrak{B}$.

a finite number of elements $a_i \in \mathfrak{B}$ $(i = 1, 2, \cdots, n)$ such that $f \in E_{a_i}$ $(i = 1, 2, \cdots, n)$ and $\Omega - U(f) \subseteq \sum_{i=1}^{n} (\Omega - E_{a_i})$ i.e., $\prod_{i=1}^{n} E_{a_i} \subseteq U(f)$. If we now put $a_0 = a_1 \wedge a_2 \wedge \cdots \wedge a_n$, then we have $f \in E_{a_0} = \prod_{i=1}^{n} E_{a_i} \subseteq U(f)$. This proves the total disconnectedness of Ω.

Lastly, we shall prove that for any open-and-closed subset E of Ω, there exists an element $a_0 \in \mathfrak{B}$ such that $E = E_{a_0}$. For any $f \in E$, there exists, by the above result, an $a \in \mathfrak{B}$ such that $f \in E_a \subseteq E$. Since each E_a is open and since E is compact, there exists a finite number of elements $a_i \in \mathfrak{B}$ $(i = 1, 2, \cdots, n)$ such that $E = \sum_{i=1}^{n} E_{a_i}$. If we put $a_0 = a_1 \vee a_2 \vee \cdots \vee a_n$, then we have $E_{a_0} = \sum_{i=1}^{n} E_{a_i} = E$.

The proof of Theorem 12 is hereby completed.

15. Case of discrete spaces 2. Banach limits

Let (m) be a space of all bounded sequences of real numbers $x = \{x_n\}$, $n = 1, 2, \cdots$ with $\| x \| = \sup_n | x_n |$ as its norm. A functional Lim x_n defined on (m) is called a Banach limit, if it satisfies the following conditions.

$$\text{Lim}\ (x_n + y_n) = \text{Lim}\ x_n + \text{Lim}\ y_n, \tag{15.1}$$

$$\text{Lim}\ (\lambda x_n) = \lambda\ \text{Lim}\ x_n, \tag{15.2}$$

$$x_n \geqq 0, \quad n = 1, 2, \cdots \quad \textit{imply} \quad \text{Lim}\ x_n \geqq 0, \tag{15.3}$$

$$x_n = 1, \quad n = 1, 2, \cdots \quad \textit{imply} \quad \text{Lim}\ x_n = 1, \tag{15.4}$$

$$\text{Lim}\ x_{n+1} = \text{Lim}\ x_n. \tag{15.5}$$

If we consider a discrete space Ω' consisting of a denumerable number of isolated points $\{t_n\}$ $(n = 1, 2, \cdots)$, then (m) is nothing but a space $C(\Omega')$ of all bounded continuous real-valued functions $x(t)$ defined on Ω' $(x_n = x(t_n), n = 1, 2, \cdots)$. If we now consider $C(\Omega')$ as an abstract (M)-space in the usual manner, then $C(\Omega')$ is isometric and lattice isomorphic to the space $C(\Omega)$, where $\Omega = \Omega(\aleph_0)$ is a compact space which was defined in the beginning of section 14. If we put $f(x) = \text{Lim}\ x_n$, then the functional $f(x)$ defined on $C(\Omega) = C(\Omega')$ satisfies the following conditions:

$$f(x + y) = f(x) + f(y), \tag{15.6}$$

$$f(\lambda x) = \lambda f(x), \tag{15.7}$$

$$x \geqq 0 \quad \textit{implies} \quad f(x) \geqq 0, \tag{15.8}$$

$$f(\mathbf{1}) = 1, \quad \textit{where} \quad \mathbf{1} = (1, 1, \cdots), \tag{15.9}$$

$$f(T(x)) = f(x), \quad \textit{where} \quad T(x) = (x_2, x_3, \cdots) \text{ if } x = (x_1, x_2, \cdots). \tag{15.10}$$

Let us observe only the conditions (15.6)–(15.9) for the time being. Then these conditions mean that $f(x)$ is a positive bounded linear functional with $f(\mathbf{1}) = 1$. But this last condition is equivalent to $\| f \| = 1$ (see Lemma 11.1). Consequently, by Theorem 9, there exists a completely additive, regular non-

negative set function $\mu(E)$ defined for all Borel sets E of the space Ω with $\mu(\Omega) = 1$ such that

$$f(x) = \int_\Omega x(t)\mu(dt) \tag{15.11}$$

for any $x(t) \in C(\Omega) = C(\Omega')$. Moreover, if we require an additional condition:

$$x \wedge y = 0 \quad \textit{implies} \quad f(x) = 0 \quad \text{or} \quad f(y) = 0, \tag{15.12}$$

then the measure $\mu(E)$ is concentrated in one point $t_0 \in \Omega$. If $t_0 \in \Omega - \Omega'$, then the functional $f(x) = x(t_0)$ is not trivial. It is easy to see that in this case we have

$$f(x) = 0 \textit{ for any } x \textit{ of the form } x = (x_1, x_2, \cdots, x_n, 0, 0, \cdots). \tag{15.13}$$

REMARK. It will be easily seen that under the conditions (15.6)–(15.9) the condition (15.12) is equivalent to

$$(f(x))^2 = f(x^2), \tag{15.14}$$

where $x^2 = (x_1^2, x_2^2, \cdots, x_n^2, \cdots)$ if $x = (x_1, x_2, \cdots, x_n, \cdots)$. Such functionals $f(x)$ were discussed by J. von Neumann [14], and it is to be remarked that (15.14) is equivalent to

$$f(x)f(y) = f(x \cdot y), \tag{15.15}$$

where $xy = (x_1y_1, x_2y_2, \cdots, x_ny_n, \cdots)$ if $x = (x_1, x_2, \cdots, x_n, \cdots)$ and $y = (y_1, y_2, \cdots, y_n, \cdots)$. This last condition together with (15.6), (15.7) means that $x \to f(x)$ is a homomorphic mapping of a ring (m) onto a ring of real numbers.

Next we shall discuss the condition (15.10). Let us denote by Δ_μ the set of all functionals $f_\mu(x)$ of the form (15.11) with $\mu(E) \geqq 0$, $\mu(\Omega) = 1$, i.e., the functionals which satisfy (15.6)–(15.9). Then Δ_μ is a compact convex set with respect to the weak topology. If we put $\tilde{f}(x) = f(T(x))$, then $f \to \tilde{f} = \varphi(f)$ is an *affine* continuous mapping of Δ_μ into itself, where under affine mapping we mean a mapping φ such that $\varphi(\lambda f + \lambda' g) = \lambda\varphi(f) + \lambda'\varphi(g)$ for any $f, g \in \Delta_\mu$; $\lambda, \lambda' \geqq 0$, $\lambda + \lambda' = 1$. It is easy to see that the condition (15.10) means that f is a fixed point under this affine mapping φ of Δ_μ into itself, and the existence of such a fixed point follows from the following

LEMMA 15.1. *Let E be a Banach space, and let Δ be a convex compact set in the conjugate space E^* of E (with respect to the weak topology). If $\varphi(f)$ is a continuous affine mapping of Δ into itself, then there exists a fixed element $f_0 \in \Delta$ such that $\varphi(f_0) = f_0$.*

PROOF. Let Δ_n be the set of all elements of the form: $n^{-1}(f + \varphi(f) + \cdots + \varphi^{n-1}(f))$, $f \in \Delta$. Then Δ_n is a non-empty compact subset of Δ and the family $\{\Delta_n\}$ $(n = 1, 2, \cdots)$ has clearly a finite intersection property: for any n_i $(i = 1, 2, \cdots, k)$ $\Delta_{n_1} \cdot \Delta_{n_2} \cdot \cdots \cdot \Delta_{n_k} \neq 0$. This follows directly from $\Delta_{n_1} \cdot \Delta_{n_2} \cdot \cdots \cdot \Delta_{n_k} \supseteq \Delta_{n_1 n_2 \cdots n_k} \neq 0$. Hence, there exists an $f_0(x)$ which belongs to all Δ_n,

$n = 1, 2, \cdots$. It is easy to see that this f_0 is a required one. (For generalizations of this lemma, see [8].)

REMARK. We can also prove the existence of a fixed point in the folllowing way: Let $S(x)$ be a mapping defined on (m) by $x \to S(x) = x'$: $x = (x_1, x_2, \cdots, x_n, \cdots)$, $x' = (x_1, 2^{-1}(x_1 + x_2), \cdots, n^{-1}(x_1 + x_2 + \cdots + x_n), \cdots)$, and put $f_0(x) = f(S(x))$, where $f(x)$ is an arbitrary bounded linear functional which satisfies (15.6)–(15.9) and (15.13). Then we have $f_0(T(x)) = f_0(x)$. Indeed, writing $f(x) = f(x_1, x_2, \cdots, x_n, \cdots)$ for $x = (x_1, x_2, \cdots) \in (m)$, we have $f_0(x) - f_0(T(x)) = f(S(x)) - f(S(T(x))) = f(x_1, 2^{-1}(x_1 + x_2), \cdots, n^{-1}(x_1 + x_2 + \cdots + x_n) \cdots) - f(x_2, 2^{-1}(x_2 + x_3), \cdots, n^{-1}(x_2 + x_3 + \cdots + x_{n+1}), \cdots) = f(0, 0, \cdots, 0, (n+1)^{-1}(x_1 + x_2 + \cdots + x_{n+1}), (n+2)^{-1}(x_1 + x_2 + \cdots + x_{n+2}), \cdots) - f(0, 0, \cdots, 0, (n+1)^{-1}(x_2 + x_3 + \cdots + x_{n+2}), (n+2)^{-1}(x_2 + x_3 + \cdots + x_{n+3}), \cdots) = f(0, 0, \cdots, 0, (n+1)^{-1}(x_1 - x_{n+2}), (n+2)^{-1}(x_1 - x_{n+3}), \cdots)$. Hence, $| f_0(x) - f_0(T(x)) | \leqq f(0, 0, \cdots, 0, 2(n+1)^{-1} \| x \|, 2(n+2)^{-1} \| x \|, \cdots) \leqq 2n^{-1} \| x \|$ for $n = 1, 2, \cdots$. Since n is arbitrary, we must have $f_0(x) = f_0(T(x))$.

Summing up the results, we have

THEOREM 13. *Every Banach limit which satisfies* (15.1)–(15.5) *can be considered as a bounded linear functional* $f(x)$ *defined on a Banach space* $C(\Omega(\aleph_0))$ *which satisfies the conditions* (15.6)–(15.10). (*For the definition of* $\Omega(\aleph_0)$ *see the beginning of section* 14.) *The functionals* $f(x)$ *which satisfy* (15.6)–(15.9) *can be expressed as an integral* (15.11). *The totality of all such functionals forms a compact convex set* Δ_μ *in the conjugate space* $C^*(\Omega(\aleph_0))$ *of* $C(\Omega(\aleph_0))$, *and every Banach limit* (*which satisfies the additional condition* (15.10)) *can be obtained as a fixed point of the affine mapping* φ *of* Δ_μ *into itself.*

It is easy to see that the result of this section can be extended to the case where the power of a given discrete space Ω' is greater than $\aleph_0$.

16. Relations with abstract (L)-spaces

A Banach space (AL) is called an *abstract* (L)*-space* if it is a linear lattice, and if it satisfies, besides the axioms (1.1)–(1.8), the following condition:

(16.1) $\quad x \geqq 0, \quad y \geqq 0$ *imply* $\| x + y \| = \| x \| + \| y \|$.

It is easy to see that the condition (1.16) is also satisfied. Such a Banach space was introduced by G. Birkhoff [1], and it was shown that a large number of results concerning dependent probabilities can be extended to such spaces.

THEOREM 14. *For any abstract* (L)*-space* (AL) *with an F-unit* (*see section* 3), *there exists a totally disconnected compact Hausdorff space* Ω *and a completely additive measure* $m(E)$ *defined for all Borel sets* E *of* Ω *such that* (AL) *is isometric and lattice isomorphic to the space* $L(\Omega; m)$ *of all measurable functions* $x(t)$ *which are integrable with respect to* $m(E)$ *on* Ω $\left(\| x \| = \int_\Omega | x(t) | \, m(dt); x \geqq y: x(t) \geqq y(t) \right.$ *almost everywhere on* $\left. \Omega \right)$.

This theorem was proved in [9], [11].

REMARK 1. In my former paper [11] it was shown that $m(E)$ is defined on the smallest Borel field which contains all open-and-closed subsets of Ω. It is, however, easy to see that this measure can be extended to a completely additive measure $m'(E)$ defined for all Borel sets E of Ω. Moreover, this extension is unique if we require that $m'(E)$ is regular, and it is easy to see that we have $L(\Omega; m) = L(\Omega; m')$. This follows from the fact that, for any Borel set E and for any $\epsilon > 0$, there exists an open-and-closed set E' such that $m'(E + E' - E \cdot E') < \epsilon$. (See Remark 2, after Theorem 9.)

REMARK 2. Theorem 14 is true in the following sense even if there exists no F-unit in (AL). In this case, (AL) is decomposed into a direct sum of abstract (L)-spaces $(AL)_\alpha$ with an F-unit ($\alpha \in \mathfrak{M}$, where $\mathfrak{M}$ is a non-denumerable set of indices α), and each $(AL)_\alpha$ is concretely represented as in Theorem 14. Hence there exists a family of compact Hausdorff spaces Ω_α, $\alpha \in \mathfrak{M}$ such that $(AL) = \sum_{\alpha \in \mathfrak{M}} (AL)_\alpha$ is isometric and lattice isomorphic to $L(\Omega; m) = \sum_{\alpha \in \mathfrak{M}} L(\Omega_\alpha; m)$, where $\Omega = \sum_{\alpha \in \mathfrak{M}} \Omega_\alpha$ and each element of $L(\Omega; m)$ is a measurable function $x(t)$ defined on $\Omega = \sum_{\alpha \in \mathfrak{M}} \Omega_\alpha$ such that there exists a sequence of indices α_n, $n = 1, 2, \cdots$ such that $\| x \| = \sum_{n=1}^{\infty} \int_{\Omega_{\alpha_n}} | x(t) | m(dt) < \infty$ and $x(t) = 0$ almost everywhere on Ω_α, $\alpha \neq \alpha_n$, $n = 1, 2, \cdots$. (For details, see [11].)

We shall discuss in this chapter the relations between abstract (M)-spaces and abstract (L)-spaces. We begin with

LEMMA 16.1. *Let E be a Banach lattice, i.e., a Banach space with a partial ordering relation $x \geqq y$ which satisfies the conditions* (1.1)–(1.8) *and* (1.16). *Then the conjugate space E^* of E is also a Banach lattice, if we put $f \geqq g$ if and only if $f(x) \geqq g(x)$ for any $x \geqq 0$, $x \in E$.*

PROOF. We have only to show that the conditions (1.5), (1.8) and (1.16) are satisfied for the partial ordering $f \geqq g$. To define $h = f \vee g$, put $h(x) =$ l.u.b. $\{f(x_1) + g(x_2)\}$ for $x \geqq 0$, where l.u.b. means the least upper bound for all decompositions $x = x_1 + x_2$, $x_1 \geqq 0$, $x_2 \geqq 0$. $h(x)$ is thus defined for $x \geqq 0$, and if we put $h(x) = h(x \vee 0) - h((-x) \vee 0)$, then this $h(x)$ is a required one (see, for example, H. Freudenthal [5]).

Next, to show that $f \geqq g \geqq 0$ implies $\| f \| \geqq \| g \| \geqq 0$, let $\epsilon > 0$ be an arbitrary positive number. Then there exists an $x \in E$, such that $\| x \| \leqq 1$ and $g(x) > \| g \| - \epsilon$. We may assume that $x \geqq 0$, since otherwise we can replace x by $x \vee 0$. Indeed, we have $\| x \vee 0 \| \leqq \| x \vee 0 + (-x) \vee 0 \| = \| x \| \leqq 1$ and $g(x \vee 0) \geqq g(x) > \| g \| - \epsilon$. Consequently, we have $f(x) \geqq g(x) > \| g \| - \epsilon$ for an x with $\| x \| \leqq 1$, and this implies $\| f \| \geqq \| g \| - \epsilon$. Since $\epsilon > 0$ is arbitrary, we have $\| f \| \geqq \| g \|$.

Finally, to prove that $f \wedge g = 0$ implies $\| f + g \| = \| f - g \|$, let $\epsilon > 0$ be an arbitrary positive number. Then there exists an $x \in E$ such that $\| x \| \leqq 1$ and $f(x) + g(x) > \| f + g \| - \epsilon$. We may again assume that $x \geqq 0$. Since $f \wedge g = 0$, there exists a decomposition $x = x_1 + x_2$, $x_1 \geqq 0$, $x_2 \geqq 0$ such that

$0 \leqq f(x_1) + g(x_2) < \epsilon$.[14] If we now put $x' = -x_1 + x_2$, then we have $\| x' \| = \| (x_1 - x_2) \vee 0 + (x_2 - x_1) \vee 0 \| \leqq \| x_1 + x_2 \| = \| x \| \leqq 1$ and $f(x') - g(x') = -f(x_1) + f(x_2) + g(x_1) - g(x_2) > f(x_1) + f(x_2) + g(x_1) + g(x_2) - 2\epsilon = f(x) + g(x) - 2\epsilon > \| f + g \| - 3\epsilon$. Hence we have $\| f - g \| > \| f + g \| - 3\epsilon$, and since $\epsilon > 0$ is arbitrary, we must have $\| f - g \| \geqq \| f + g \|$.

In order to prove the inverse inequality, let $\epsilon > 0$ be an arbitrary positive number. Then there exists an $x \in E$ such that $\| x \| \leqq 1$ and $f(x) - g(x) > \| f - g \| - \epsilon$. Clearly $x' = x \vee 0 + (-x) \vee 0$ satisfies $\| x' \| = \| x \| \leqq 1$ (by (1.15)) and $f(x') + g(x') = f(x \vee 0) + f((-x) \vee 0) + g(x \vee 0) + g((-x) \vee 0) \geqq f(x \vee 0) - f((-x) \vee 0) - g(x \vee 0) + g((-x) \vee 0) = f(x) - g(x) > \| f - g \| - \epsilon$. Hence we have $\| f + g \| > \| f - g \| - \epsilon$, and since $\epsilon > 0$ is arbitrary, we must have $\| f + g \| \geqq \| f - g \|$.

THEOREM 15. *The conjugate space $(AM)^*$ of an abstract (M)-space (AM) is an abstract (L)-space, and conversely the conjugate space $(AL)^*$ of an abstract (L)-space (AL) is an abstract (M)-space.*

PROOF. By Lemma 16.1, we need only prove the following two facts:

$$(16.2)\quad f, g \in (AM)^*,\ f \geqq 0,\ g \geqq 0 \quad \textit{imply} \quad \| f + g \| = \| f \| + \| g \|,$$

$$(16.3)\quad f, g \in (AL)^*,\ f \geqq 0,\ g \geqq 0 \quad \textit{imply} \quad \| f \vee g \| = \max (\| f \|, \| g \|).$$

To prove (16.2), let $f, g \in (AM)^*$, $f \geqq 0$, $g \geqq 0$, and let $\epsilon > 0$ be an arbitrary positive number. Then there exist $x, y \in (AM)$ such that $\| x \| \leqq 1$, $\| y \| \leqq 1$, $f(x) > \| f \| - \epsilon$ and $g(y) > \| g \| - \epsilon$. We may assume $x \geqq 0$, $y \geqq 0$, since otherwise we can replace x and y by $x \vee 0$ and $y \vee 0$. Then $z = x \vee y$ satisfies $\| z \| = \max (\| x \|, \| y \|) \leqq 1$ and $f(z) + g(z) \geqq f(x) + g(y) > \| f \| + \| g \| - 2\epsilon$. Hence $\| f + g \| > \| f \| + \| g \| - 2\epsilon$, and since $\epsilon > 0$ is arbitrary, we have $\| f + g \| \geqq \| f \| + \| g \|$. The inverse inequality $\| f + g \| \leqq \| f \| + \| g \|$ is clear.

Next, in order to prove (16.3), let us first prove that there exists a unit element f_0 (in the sense of section 3) in $(AL)^*$. Indeed, if we put

$$(16.4)\qquad f_0(x) = \| x \vee 0 \| - \| (-x) \vee 0 \|$$

for any $x \in (AL)$, then we have

$$(16.5)\qquad | f_0(x) | \leqq \| x \|,$$

$$(16.6)\qquad f_0(x + y) = f_0(x) + f_0(y),$$

$$(16.7)\qquad f_0(\lambda x) = \lambda f_0(x),$$

$$(16.8)\qquad f \in (AL)^*, \quad \| f \| \leqq 1 \quad \textit{imply} \quad f \leqq f_0.$$

---

[14] $k = f \wedge g$ is defined, for $x \geqq 0$, by $k(x) = \text{g.l.b.}\ \{f(x_1) + g(x_2)\}$, where g.l.b. means the greatest lower bound for all decompositions of x: $x = x_1 + x_2$, $x_1 \geqq 0$, $x_2 \geqq 0$.

(16.5) and (16.7) are clear; (16.6) follows from the relation

$$\|(x+y)\vee 0\| + \|(-x)\vee 0\| + \|(-y)\vee 0\| = \|(-x-y)\vee 0\| + \|x\vee 0\| + \|y\vee 0\|,$$

which is a direct consequence of (16.1) and the equality

$$(x+y)\vee 0 + (-x)\vee 0 + (-y)\vee 0 = (-x-y)\vee 0 + x\vee 0 + y\vee 0.$$

(16.8) is also clear, since $\|f\| \leqq 1$, $x \geqq 0$ imply $f(x) \leqq \|x\| = \|x\vee 0\| - \|(-x)\vee 0\| = f_0(x)$.

Thus we have proved that $f_0(x)$ is a unit element of $(AL)^*$ in the sense of section 3. Hence, by Remark 1 of section 3, $(AL)^*$ is an abstract (M)-space. The proof of Theorem 15 is completed.

EXAMPLE 1. If $(AM) = C(\Omega)$, where Ω is a compact Hausdorff space, then $(AM)^* = \mathfrak{M}(\Omega)$ (see Theorem 10). $\mathfrak{M}(\Omega)$ is clearly an abstract (L)-space. For further examples, see Remark after Theorem 10.

EXAMPLE 2. If $(AL) = L(\Omega; m)$, where Ω is a compact Hausdorff space with a completely additive measure $m(E)$ such that $m(\Omega) = 1$ (see Theorem 14), then $(AL)^* = M(\Omega; m)$ (for the definition of $M(\Omega; m)$ see section 2). The same is also true if $(AL) = L(\Omega; m) = \sum_{\alpha\in\mathfrak{M}} L(\Omega_\alpha; m)$, where $\Omega = \sum_{\alpha\in\mathfrak{M}} \Omega_\alpha$ and each Ω_α is compact.

REMARK 1. Theorem 15 does not mean that (AM) and (AL) are reflexive. In fact, these are reflexive only when they are finite dimensional.

REMARK 2. For any abstract (M)-space (AM), its second conjugate space $(AM)^{**}$ is also an abstract (M)-space (by using Theorem 15 twice) which contains (AM) as a closed linear subspace.[15] The same argument is true also for

---

[15] We should be careful about the following situation: (AM) is embedded isometrically into $(AM)^{**}$ and this embedding is partial order preserving, i.e. if $x, y \in (AM)$ correspond to $X, Y \in (AM)^{**}$, then $x \geqq y$ implies $X \geqq Y$. But it is not obvious that in this case $x \vee y$ corresponds to $X \vee Y$. In order to show this, let us put $z = x \vee y$, $Z = X \vee Y$, and let Z' be the element of $(AM)^{**}$ to which z corresponds. It is then clear that we have $Z' \geqq Z$ (this is a direct consequence of $Z' \geqq X$ and $Z' \geqq Y$). In order to prove the inverse inequality, let us remember that, for $f \geqq 0$, $Z(f)$ is defined by $Z(f) = \text{l.u.b.}\{X(f_1) + Y(f_2)\} = \text{l.u.b.}\{f_1(x) + f_2(y)\}$, where l.u.b. means the least upper bound for all decompositions of f: $f = f_1 + f_2$, $f_1 \geqq 0, f_2 \geqq 0$. On the other hand, if we represent (AM) concretely as a subspace of the space $C(\Omega)$ of all real valued continuous functions defined on a compact Hausdorff space Ω (Theorem 1), then $z(t) = \max(x(t), y(t))$ for all $t \in \Omega$, and every $f \in (AM)^*$ with $f \geqq 0$ may be considered as a regular measure $\mu(E)$ defined for all Borel sets E of Ω (Theorem 9). Let $O \subseteq \Omega$ be the set of all t at which $x(t) > y(t)$. Then O is an open set which is at the same time an F_σ. Let us now define two measures $\mu_1(E)$ and $\mu_2(E)$ by $\mu_1(E) = \mu(EO)$ and $\mu_2(E) = \mu(E - EO)$. Then $\mu_1(E)$ and $\mu_2(E)$ are both regular measures defined on Borel sets of Ω, and if we denote by f_1 and f_2 the corresponding elements of $(AM)^*$, then we have $f = f_1 + f_2$, $f_1 \geqq 0, f_2 \geqq 0$ and it is not difficult to see that $f(z) = f_1(x) + f_2(y)$. Hence $Z(f) \geqq f(z) = Z'(f)$ for every $f \geqq 0$, i.e. $Z \geqq Z'$. Thus we proved the inverse inequality, and hence $Z = Z'$.

(AL) and $(AL)^{**}$.[16] It is, however, to be noted that $(AM)^{**}$ is always an abstract (M)-space with a unit element, while this is not the case for $(AL)^{**}$ even if (AL) has an F-unit.

Hence, every abstract (M)-space can be embedded in an abstract (M)-space with a unit element. Consequently, in discussing the realization of abstract (M)-spaces, we have only to treat the case when there exists a unit element. This fact explains the relation between Theorem 1 and Theorem 2.

REMARK 3. Let (AL) be an abstract (L)-space. Then $(AL)^*$ is an abstract (M)-space with a unit element. Hence, by Theorem 2, there exists a compact Hausdorff space Ω such that $(AL)^*$ is isometric and lattice isomorphic to $C(\Omega)$. Hence every element $x \in (AL)$ can be considered as a bounded linear functional $x(f)$ defined on $C(\Omega)$. Consequently, by Theorem 10, $x(f)$ can be represented as an integral:

$$x(f) = \int_\Omega f(t)\mu(dt) \tag{16.9}$$

where $\mu(E)$ is a completely additive regular real-valued set function defined for all Borel sets E of Ω. From this we have

THEOREM 16. *For any abstract (L)-space (AL), there exists a compact Hausdorff space Ω such that (AL) is isometric and lattice isomorphic to a closed linear subspace of the space $\mathfrak{M}(\Omega)$ of all completely additive regular real-valued set-functions $\mu(E)$ defined for all Borel sets E of Ω (where norm and partial ordering are defined as in Theorem 10).*

In the same manner we can prove (see Theorem 14, Remark 2 after Theorem 14, and Example 2 after Theorem 15) the following

THEOREM 17. *For any abstract (M)-space (AM), there exists a Hausdorff space $\Omega = \sum_{\alpha\in\mathfrak{M}} \Omega_\alpha$, where each Ω_α is compact, and a completely additive measure defined for all Borel sets of each Ω_α such that (AM) is isometric and lattice isomorphic to a closed linear subspace of the space $M(\Omega; m) = \sum_{\alpha\in\mathfrak{M}} M(\Omega_\alpha; \mathfrak{M})$ of all bounded measurable functions $x(t)$ defined on Ω, where we call $x(t)$ measurable if and only if $x(t)$ is measurable on each Ω_α. (For the definition of $M(\Omega; m)$ see section 2.)*

Theorems 16 and 17 are weaker than Theorems 14 and 1 (or 2) respectively.

INSTITUTE FOR ADVANCED STUDY

BIBLIOGRAPHY

(1) G. Birkhoff, *Dependent probabilities and the space (L)*, Proc. Nat. Acad. Sci., U.S.A., **24** (1938), 154–159.

[16] It is also not difficult to see that the embedding of (AL) into $(AL)^{**}$ is not only isometric and partial order preserving but also lattice isomorphic. We have only to consider the concrete representation of (AL) and the general form of positive linear functionals defined on this concrete (AL) (see Example 2 above). The arguments of the same kind as in footnote 15 will show that, if $x, y \in (AL)$ correspond to $X, Y \in (AL)^{**}$, then $z = x \vee y$ corresponds to $Z = X \vee Y$. Recently it was proved by Dr. Alaoglu that when we embed an arbitrary Banach lattice E into its second conjugate E^{**} (in the ordinary way) this embedding is always a lattice isomorphism.

(2) G. Birkhoff, *Lattice theory*, Amer. Math. Soc. Colloquium Publications, 1940.
(3) H. F. Bohnenblust, Characterization of L_p-spaces, Duke Math. Jour., **6** (1940), 627–640.
(4) E. Čech, *On bicompact spaces*, Annals of Math., **38** (1937), 823–844.
(5) H. Freudenthal, *Teilweise geordnete Moduln*, Proc. Acad. Amsterdam, **39** (1936), 641–651.
(6) I. Gelfand, *On normed rings*, C. R. URSS, **23** (1939), 430–432.
(7) I. Gelfand and A. Kolmogoroff, *On rings of continuous functions of topological spaces*, C. R. URSS, **22** (1939), 11–15.
(8) S. Kakutani, *Two fixed point theorems concerning bicompact convex sets*, Proc. Imp. Acad. Japan, **14** (1938), 242–245.
(9) S. Kakutani, *Mean ergodic theorem in abstract (L)-spaces*, Proc. Imp. Acad. Japan, **15** (1939), 121–123.
(10) S. Kakutani, *Weak topology, bicompact sets and the principle of duality*, Proc. Imp. Acad. Japan, **16** (1940), 63–67.
(11) S. Kakutani, *Concrete representation of abstract (L)-spaces and the mean ergodic theorem*, Annals of Math., **42** (1941), 523–537.
(12) L. Kantorovitch, *Lineare halbgeordnete Räume*, Recueil Math., **44** (1937), 121–165.
(13) J. von Neumann, *On an algebraic generalization of quantum mechanical formalism*, Part 1, Recueil Math., **43** (1936), 415–484.
(14) J. von Neumann, *Invariant measure*, Princeton lectures, 1940–41.
(15) S. Saks, *Integration in abstract metric spaces*, Duke Math. Jour., **4** (1938), 408–411.
(16) G. Šilov, *Ideals and subrings of the ring of continuous functions*, C. R. URSS, **22** (1939), 7–10.
(17) M. H. Stone, *Applications of the theory of Boolean rings to general topology*, Trans. Amer. Math. Soc., **41** (1937), 376–381.
(18) M. H. Stone, *A general theory of spectra*, Proc. Nat. Acad. Sci., U. S. A., **26** (1940), 280–283.
(19) K. Yosida, *On the theory of spectra*, Proc. Imp. Acad. Japan, **16** (1940) 578–583.

ANNALS OF MATHEMATICS
Vol. 42, No. 4, October, 1941

CONCRETE REPRESENTATION OF (*M*)-SPACES

BY H. F. BOHNENBLUST AND S. KAKUTANI

(Received January 20, 1941)

1. Introduction

In the preceding paper[1] one of us proves that any Banach lattice which satisfies the condition

(1.1) $\quad 0 < x, \quad 0 < y$ implies $\| x \vee y \| = \text{Max}\,(\| x \|, \| y \|)$

can be represented as a sublattice of the lattice of all real valued continuous functions over a compact[2] Hausdorff space Ω. The purpose of this note is to obtain the same result from the weaker condition

(1.2) $\quad x \wedge y = 0$ implies $\| x \vee y \| = \text{Max}\,(\| x \|, \| y \|)$.

Our proof yields, in particular, the result that in a Banach lattice (1, 2) implies (1, 1). It does not seem possible, however, to establish this implication directly. Our method consists in obtaining the representation of the lattice independently of the previous result and then (1.1) is trivially true for continuous functions.

If the additional assumption is made that the Banach lattice contains a unit element, it is shown in the preceding paper that the representation uses *all* continuous functions. In this form the result was obtained independently by Mark and Selim Krein.[3] The method we shall use in the present paper is an adaptation of that followed in Krein's paper.

2

Let E be a Banach lattice, i.e.

(2.1) $\quad E$ is a Banach space,

(2.2) $\quad E$ is a vector lattice,

(2.3) $\quad 0 < x < y$ implies $\| x \| \leqq \| y \|$,[4]

(2.4) $\quad x \wedge y = 0$ implies $\| x - y \| = \| x + y \|$.

---

[1] S. Kakutani, *Concrete representation of abstract* (*M*)-spaces. Annals of Math., 42 (1941), pp. 994–1024.

[2] We use the term "compact" as "open coverings can be reduced to finite coverings."

[3] M. and S. Krein, *On an inner characteristic of the set of all continuous functions on a bicompact Hausdorff space*, Comptes Rendus U.R.S.S., 27 (1940), 427–430.

[4] It is interesting to notice that condition (2, 3) cannot be deduced from (2.4) and (2.5) combined.

Assume furthermore that the condition

$$x \wedge y = 0 \text{ implies } \| x + y \| = \text{Max } (\| x \|, \| y \|) \tag{2.5}$$

holds in E.

The conjugate space E^* is also a Banach lattice. Denote by S^* its unit sphere ($\| f \| \leqq 1$) and by S^*_+ the positive part ($f \geqq 0$) of S^*. The set S^*_+ is *bounded, convex* and *w*-closed.*[5] The extreme points of S^*_+ form a set which will be denoted by T. By a theorem of Krein and Milman[6] we have

THEOREM 2.1. *The set S^*_+ is the smallest convex, w*-closed set which contains T.*

THEOREM 2.2. *With the exception of 0, every element of T has norm 1.*

PROOF: If $f \neq 0$ belongs to T, then $f/\| f \|$ belongs to S^*_+ and the element f is an *interior* point of the segment joining 0 and $f/\| f \|$ unless $\| f \| = 1$.

THEOREM 2.3. *If $f \in T$ and if $u, v \in E$, $u \wedge v = 0$ then $f(u) \cdot f(v) = 0$.*

The proof is by contradiction. Assume $f \in T, f(u) > 0, f(v) > 0$ and $u \wedge v = 0$. Two elements of g, $h \in E^*$ will be constructed such that

$$\begin{aligned} &\text{i)}\ \ 0 \leqq g \leqq f \qquad &&\text{ii)}\ \ g(v) = 0 \\ &\text{iii)}\ \ g(u) = f(u) > 0 \qquad &&\text{iv)}\ \ f = g + h \\ &\text{v)}\ \ 1 = \| f \| = \| g \| + \| h \|. \end{aligned} \tag{2.6}$$

Conditions i) and iii) show that $0 < h$. Thus $g/\| g \|$ and $h/\| h \|$ belong to S^*_+ and

$$f = \| g \| \cdot \frac{g}{\| g \|} + h \cdot \frac{h}{\| h \|}$$

proves that f is an interior point of the segment joining $g/\| g \|$ and $h/\| h \|$. This is a contradiction to the assumption $f \in T$ since $g/\| g \| \neq h/\| h \|$ by ii) and iii).

Definition of g. For $x \geqq 0$ put

$$g(x) = \sup_n f(nu \wedge x)$$

and for $x = x^+ - x^-$, $g(x) = g(x^+) - g(x^-)$. We omit the details showing that $g \in E^*$ and that i), ii), iii) are satisfied.

Definition of h: $h = f - g$. Condition iv) is true and to prove the last one, v), we need only show that

$$\| g \| + \| h \| \leqq 1 + 4\epsilon$$

for any $\epsilon > 0$. Choose $\epsilon > 0$ and determine x_1 such that $\| x_1 \| = 1$ and $g(x_1) > \| g \| - \epsilon$. We may assume $x_1 > 0$ since otherwise it could be replaced by $| x_1 |$.

---

[5] w^*-topology in E^* is the weak topology with reference to the elements of E.

[6] M. Krein and D. Milman, *On extreme points of regular convex sets*, Studia Math., IX (1940), 133–138.

By the definition of g there will exist a positive integer n_1 for which

$$f(n_1 u \wedge x_1) > g(x_1) - \epsilon > \| g \| - 2\epsilon.$$

Put $x_2 = x_1 \wedge n_1 \cdot u$; then

$$\| x_2 \| \leqq 1, \qquad g(x_2) = f(x_2) > \| g \| - 2\epsilon, \qquad h(x_2) = 0. \tag{2.7}$$

Now we turn our attention to h. Determine an element y_1 such that

$$0 \leqq y_1, \qquad \| y_1 \| = 1, \qquad h(y_1) > \| h \| - \epsilon,$$

and choose n_2 to guarantee $1 < \epsilon \cdot (n_2 + 1)$. Put

$$y_2 = (y_1 - n_2 \cdot x_2) \vee 0.$$

Obviously $0 \leqq y_2 \leqq y_1$ and hence $\| y_2 \| \leqq 1$. Also

$$h(y_2) = h(y_1 \vee n_2 x_2) - n_2 h(x_2) = h(y_1 \vee n_2 x_2) \geqq h(y_1) > \| h \| - \epsilon. \tag{2.8}$$

We evaluate the norm of $x_2 \wedge y_2$. Notice that

$$x_2 \wedge y_2 = 0 \vee [x_2 \wedge (y_1 - n_2 \cdot x_2)].$$

But

$$n_2[x_2 \wedge (y_1 - n_2 x_2)] \leqq n_2 x_2$$

$$[x_2 \wedge (y_1 - n_2 x_2)] \leqq y_1 - n_2 x_2$$

and thus by addition

$$(n_2 + 1)[x_2 \wedge (y_1 - n_2 x_2)] \leqq y_1.$$

This implies $0 \leqq x_2 \wedge y_2 \leqq y_1/(n_2 + 1) \leqq \epsilon \cdot y_1$ and therefore

$$\| x_2 \wedge y_2 \| \leqq \epsilon. \tag{2.9}$$

Finally let us put

$$x_3 = x_2 - (x_2 \wedge y_2) \qquad \text{and} \qquad y_3 = y_2 - (x_2 \wedge y_2).$$

Evidently $x_3 \wedge y_3 = 0$, $\| x_3 \| \leqq 1$, $\| y_3 \| \leqq 1$, $\| x_3 + y_3 \| = \text{Max}\,(\| x_3 \|, \| y_3 \|) \leqq 1$. Since $0 \leqq h(x_2 \wedge y_2) \leqq h(x_2) = 0$ it follows from (2.8) that

$$h(y_3) = h(y_2) > \| h \| - \epsilon.$$

On the other hand, by (2.7) and (2.9)

$$g(x_3) = g(x_2) - g(x_2 \wedge y_2) > \| g \| - 2\epsilon - \| g \| \cdot \epsilon \geqq \| g \| - 3\epsilon.$$

Combining these results we obtain

$$1 \geqq f(x_3 + y_3) = g(x_3) + g(y_3) + h(x_3) + h(y_3)$$

$$> \| g \| - 3 \cdot \epsilon + 0 + 0 + \| h \| - \epsilon = \| g \| + \| h \| - 4\epsilon.$$

The proof of v) is complete and theorem 2.3 is established.

THEOREM 2.4. *Let Ω be the w^*-closure of T. If $f \in \Omega$ then $f \geqq 0$ and if u, $v \in E$, $u \wedge v = 0$, then $f(u) \cdot f(v) = 0$.*

PROOF: For any $\epsilon > 0$ there exists $g \in T$ with

$$| g(u) - f(u) | < \epsilon,$$

$$| g(v) - f(v) | < \epsilon.$$

By theorem 2.3 either $g(u)$ or $g(v) = 0$. Hence either $|f(u)|$ or $|f(v)| < \epsilon$, which implies of course $f(u) \cdot f(v) = 0$. The proof of $f \geqq 0$ is even simpler.

3. The Representation

Each element x of E determines a function φ_x defined over Ω by $\varphi_x(f) = f(x)$. The set Ω will be considered as a topological space by taking the w^*-topology of E^*. It is a Hausdorff space and since S^* is w^*-compact, Ω will also be w^*-compact. In the space Ω the function φ_x is continuous. Furthermore we have evidently

$$\varphi_{\lambda x + \mu y} = \lambda \varphi_x + \mu \varphi_y \tag{3.1}$$

$$x \geqq 0 \text{ implies } \varphi_x \geqq 0. \tag{3.2}$$

[The last inequality means: for any $f \in \Omega$, $\varphi_x(f) \geqq 0$].

If $x_1 \wedge x_2 = 0$ then $f(x_1) \cdot f(x_2) = 0$ for any $f \in \Omega$ by theorem 2.4. Thus $\varphi_{x_1}(f) \cdot \varphi_{x_2}(f) = 0$ and therefore $\varphi_{x_1} \wedge \varphi_{x_2} = 0$. For any x_1 and x_2, the elements $x_1 - (x_1 \wedge x_2)$ and $x_2 - (x_1 \wedge x_2)$ are disjoint. We have then

$$(\varphi_{x_1} - \varphi_{x_1 \wedge x_2}) \wedge (\varphi_{x_2} - \varphi_{x_1 \wedge x_2}) = 0$$

or

$$\varphi_{x_1} \wedge \varphi_{x_2} = \varphi_{x_1 \wedge x_2} \quad \text{and also} \quad \varphi_{x_1} \vee \varphi_{x_2} = \varphi_{x_1 \vee x_2} \tag{3.3}$$

Finally, if we define $\| \varphi_x \| = \max | \varphi_x(f) |$ $(f \in \Omega)$, we prove

$$\| x \| = \| \varphi_x \|. \tag{3.4}$$

It is sufficient to prove (3.5) for positive x, since $\| |x| \| = \| x \|$ and $\| \varphi_{|x|} \| = \| \varphi_x \|$. Evidently $\| \varphi_x \| \leqq \| x \|$. Let $f_0 \in E^*$, $\| f_0 \| = 1, f_0(x) = 1$. We may assume $f_0 \geqq 0$ since otherwise $|f_0|$ could be taken. If now $\| \varphi_x \| = \| x \| - \epsilon$, $\epsilon > 0$ then

$$f \in \Omega \to f(x) \leqq \| x \| - \epsilon,$$

thus

$$f \in S^*_+ \quad f(x) \leqq \| x \| - \epsilon$$

and we obtain a contradiction for f_0.

The equation (3.4) shows in particular that $\varphi_x = 0$ implies $x = 0$, in other words: to different x's correspond different φ_x's.

The relations (3.1), (3.2), (3.3) and (3.4) give us the desired representation.

PRINCETON, N. J.

ANNALS OF MATHEMATICS
Vol. 42, No. 4, October, 1941

SELECTIVE EQUATIONS

BY E. T. BELL

(Received October 26, 1940)

A detail in the systematic solution of certain compound multiplicative diophantine systems[1] leads to a new type of completely solvable diophantine problem which may be of independent interest, as but few such problems are known.

1. Selective operators and systems

If $n_1, \cdots, n_s$ are integers $\geqq 0$, and if $s > 1$, $(n_1, \cdots, n_s)'$ denotes the least, and $[n_1, \cdots, n_s]'$ the greatest, of $n_1, \cdots, n_s$ if at least two of these integers are unequal; if $n_1 = \cdots n_s = n$, the value of each of these symbols is n. By definition, if m is an integer $\geqq 0$, $(m)' = [m]' = m$. Hence, for $s \geqq 1$,

$$[(n_1, \cdots, n_s)']' = (n_1, \cdots, n_s)' = ((n_1, \cdots, n_s)')',$$

$$([n_1, \cdots, n_s]')' = [n_1, \cdots, n_s]' = [[n_1, \cdots, n_s]']'.$$

The symbols $()'$, $[]'$ will be called simple selective operators.

Let each symbol $\{\}'$ denote a definite one of $()'$, $[]'$. In $\{n_1, \cdots, n_s\}'$ replace n_j by $\{n_{j_1}, \cdots, n_{j_{s_j}}\}'$, where the n_{j_i} are integers $\geqq 0$. In the result repeat the process, and so on, a finite number of times. The symbol $\{\cdots\{\cdots\{\cdots\}'\cdots\}'\cdots\}'$ so obtained will be called a compound selective operator. It will be seen that compound operators may be replaced by simple operators in the type of equations to be defined.

If the results of operating on the set $n_1, \cdots, n_s$ with two selective operators are equal for all sets $n_1, \cdots, n_s$, we say that the operators are identical; otherwise, the operators are distinct. The result of operating on the ordered set $n_1, \cdots, n_s$ with the selective operator σ' will be written $\sigma'(n_1, \cdots, n_s)$.

A selective system is defined as follows. Let $\sigma_i', \cdots, \sigma_j', \cdots, \sigma_k', \cdots, \sigma_l'$ be any selective operators (simple or compound, and unrestricted with respect to distinctness). The $a, \alpha, a_i, \cdots, b, \beta, b_i, \cdots, c, \gamma, c_k, \cdots, d, \delta, d_l$ are constant integers > 0, and all the x's, $\cdots$, y's, $\cdots$, z's, $\cdots$, w's are variable integers $\geqq 0$. The $A, \cdots, B, \cdots, C, \cdots, D$ are constant integers $\geqq 0$. If any of the A's are zero, the terms involving them are suppressed, and likewise for the B's, $\cdots$, C's, $\cdots$, D's. We write

$$\sigma_i'(x_{i,1}, \cdots, x_{i,m_i}) \equiv X_i', \cdots, \sigma_j'(y_{j,1}, \cdots, y_{j,n_j}) \equiv Y_j',$$

$$\cdots \qquad \cdots \qquad \cdots$$

$$\sigma_k'(z_{k,1}, \cdots, z_{k,r_k}) \equiv Z_k', \cdots, \sigma_l'(w_{l,1}, \cdots, w_{l,s_l}) \equiv W_l'.$$

[1] E. T. Bell, Proc. Nat. Acad. Sci., 26 (1940), 462–466.

A selective system is then a system of equations of the type

$$\sum_{i=1}^{a} A_i x_i + \sum_{i=1}^{\alpha} a_i X_i' = \cdots = \sum_{j=1}^{b} B_j y_j + \sum_{j=1}^{\beta} b_j Y_j',$$

$$\cdots \qquad \cdots \qquad \cdots$$

$$\sum_{k=1}^{c} C_k z_k + \sum_{k=1}^{\gamma} c_k Z_k' = \cdots = \sum_{l=1}^{d} D_l w_l + \sum_{l=1}^{\delta} d_l W_l',$$

in which the total number of equality signs is finite, and in which at least two sets of variables in different rows have at least one variable in common, while in any one row no two sets of variables are identical, but any two sets in that row may have variables in common.

It is required to find all integer values $\geqq 0$ of the variables satisfying the system. We shall say that this problem is solved when a method is prescribed for exhibiting the required integers in any given selective system in a finite number of non-tentative steps. The solution is in terms of a certain minimum number N of parameters ranging independently over all integers $\geqq 0$. The above definition of 'solution' is inserted because for even apparently trivial systems the N necessary and sufficient for a complete solution may be so large that it is impossible to write out the solution. In fact, if N' is any given positive integer, so elementary a selective system as

$$a(x_1, x_2)' + a(y_1, y_2)' = b(x_1, z_2)' + b(w_1, w_2)'$$

can be constructed for which $N > N'$. If in this example $N' = 1{,}000{,}000$, it suffices to take $a = 1000$, $b = 1001$.

A selective system is thus one of quadruple composition, the four operations involved being addition and multiplication of non-negative integers, and $()'$, $[]'$.

2. Dual of a selective system

As might be anticipated from the definitions in §1 and a usual method for finding the G. C. D. and L. C. M. of a set of integers from the canonical decompositions (into products of powers of distinct primes) of the integers concerned, a selective system may be restated in terms of G. C. D.'s, L. C. M.'s, and products instead of sums. It will be seen that the solution of either system implies that of the other. The connection between the two types is the solution of G. C. D.-L. C. M. equations discussed in §§3, 4.

With $n_1, \cdots, n_s$ as in §1, $(n_1, \cdots, n_s)$ denotes the G. C. D., and $[n_1, \cdots, n_s]$ the L. C. M., of $n_1, \cdots, n_s$; $\{\}$ denotes a definite one of $()$, $[]$. The rest of §1, down to the selective system exhibited, may be rewritten with $\{\}$, σ in place

of $\{\}'$, σ', and X_i, Y_j, Z_k, W_l in place of X_i', Y_j', Z_k', W_l'. The system is replaced by

$$\prod_{i=1}^{a} x_i^{A_i} \cdot \prod_{i=1}^{a} X_i^{a_i} = \cdots = \prod_{j=1}^{b} y_j^{B_j} \cdot \prod_{j=1}^{\beta} Y_j^{b_j},$$

$$\cdots \qquad \cdots \qquad \cdots$$

$$\prod_{k=1}^{c} z_k^{C_k} \cdot \prod_{k=1}^{\gamma} Z_k^{c_k} = \cdots = \prod_{l=1}^{d} w_l^{D_l} \cdot \prod_{l=1}^{\delta} W_l^{d_l};$$

and the rest of §1 follows with a few obvious verbal changes. This system and that in §1 will be called duals of each other, for the following reasons.

It will be shown that the complete solution of the above system is of the form

$$\begin{aligned}
x_i &= \prod \theta_s^{x_{i,s}} (i = 1, \cdots, a), \\
x_{i,t} &= \prod \theta_s^{x_{i,t,s}} (i = 1, \cdots, \alpha; t = 1, \cdots, m_i); \\
&\cdots \qquad \cdots \qquad \cdots \\
y_j &= \prod \theta_s^{y_{j,s}} (j = 1, \cdots, b), \\
y_{j,t} &= \prod \theta_s^{y_{j,t,s}} (j = 1, \cdots, \beta; t = 1, \cdots, n_j); \\
&\cdots \qquad \cdots \qquad \cdots \\
z_k &= \prod \theta_s^{z_{k,s}} (k = 1, \cdots, c), \\
z_{k,t} &= \prod \theta_s^{z_{k,t,s}} (k = 1, \cdots, \gamma; t = 1, \cdots, r_k); \\
&\cdots \qquad \cdots \qquad \cdots \\
w_l &= \prod \theta_s^{w_{l,s}} (l = 1, \cdots, d), \\
w_{l,t} &= \prod \theta_s^{w_{l,t,s}} (l = 1, \cdots, \delta; t = 1, \cdots, s_l);
\end{aligned}$$

in which the product refers to $s = 1, \cdots, N$, the θ's are integer parameters, all exponents are constant integers $\geqq 0$ and the parameters are subject to a certain finite set of conditions, called the G. C. D. conditions,[2] of the form

$$1 = \left(\prod_{i=1}^{m} \theta_{a_i}, \cdots, \prod_{i=1}^{n} \theta_{b_i}\right) = \cdots = \left(\prod_{i=1}^{u} \theta_{c_i}, \cdots, \prod_{i=1}^{v} \theta_{d_i}\right).$$

From this, the solution of the selective system in §1 is written down by replacing all multiplications by additions, and restricting the parameters $\theta_1, \cdots, \theta_N$ to range over only integers $\geqq 0$:

$$x_i = \sum_{s=1}^{N} x_{i,s}\theta_s \qquad (i = 1, \cdots, a),$$

$$x_{i,t} = \sum_{s=1}^{N} x_{i,t,s}\theta_s \quad (i = 1, \cdots, \alpha; t = 1, \cdots, m_i);$$

[2] E. T. Bell, Amer. Journ. Math., 55 (1933), 50–66.

and similarly for the rest. The G. C. D. conditions are correspondingly replaced by

$$0 = \left(\sum_{i=1}^{m} \theta_{a_i}, \cdots, \sum_{i=1}^{n} \theta_{b_i}\right)' = \cdots = \left(\sum_{i=1}^{u} \theta_{c_i}, \cdots, \sum_{i=1}^{v} \theta_{d_i}\right)'.$$

Each of these conditions is of the form

$$0 = (A_1' + \cdots + A_f', \cdots, B_1' + \cdots + B_g')',$$

in which the A', B' are integers $\geqq 0$. From the definition of $()'$, it follows that this condition is satisfied if and only if at least one of the following sets of conditions is satisfied,

$$\begin{aligned} A_1' &= \cdots = A_f' = 0, \\ &\cdots \qquad\quad \cdots \\ B_1' &= \cdots = B_g' = 0; \end{aligned}$$

and similarly for all in the original set. The complete solution of the selective system may therefore be separated into sets of solutions, all the solutions in a particular set being obtained from the general solution by suppressing those parameters which occur in a selected subset of conditions sufficient to satisfy the complete set of conditions; the remaining parameters in the solution are then integers $\geqq 0$ subject to no conditions.

A detailed proof of the duality described is quite simple, but leads to somewhat complicated formulas. It will be sufficient to indicate how the proof may be written out in detail.

Let $\quad p_1 = 2, \quad p_2 = 3, \quad p_3 = 5, \quad p_4 = 7, \cdots$

be the natural primes in ascending order. The canonical decompositions of the positive integers n, n_i, $n_{i,j}$, $\cdots$ are written

$$n = \prod p_\xi^{n_\xi}, \qquad n_i = \prod p_\xi^{n_{i,\xi}}, \qquad n_{i,j} = \prod p_\xi^{n_{i,j,\xi}}, \cdots,$$

where the product refers to $\xi = 1, 2, 3, \cdots$. (In other connections, this device is useful for absolute constants; for example, $10 = p_1^{10_1} p_2^{10_2} p_3^{10_3} \cdots$, so that $10_1 = 1$, $10_2 = 0$, $10_3 = 1$, $10_s = 0$, $s > 3$.) In each decomposition there is a finite index ξ' such that $\xi = 0$ if $\xi \geqq \xi'$, $\xi > 0$ if $\xi < \xi'$. This notational device is applied to the variables of the system in this section, when it follows that the system is equivalent to the selective system in §1. For if $\prod p_\xi^{f_\xi} = \prod p_\xi^{g_\xi}$, then $f_\xi = g_\xi$ for all integers $\xi > 0$; also,

$$\begin{aligned} \{u_1, \cdots, u_n\}^m &= \{\prod p_\xi^{u_{1,\xi}}, \cdots, \prod p_\xi^{u_{n,\xi}}\}^m, \\ &= \prod p_\xi^{m\{u_{1,\xi}, \cdots, u_{n,\xi}\}'}, \end{aligned}$$

by the definitions of $\{\}$, $\{\}'$. The parameters θ in the solution are replaced by their canonical decompositions, $\theta_j = \prod p_\xi^{\theta_{j,\xi}}$; whence the rule for obtaining the solution of the dual selective system is immediate. It is not necessary to include

the signs of the numbers generated as the θ's in the solution of the $\{\}$-equation range over all integers; but if desired this may be done by noting that for any integer c, $c = \operatorname{sgn} c \cdot \prod p_\xi^{c_\xi}$, where sgn has its usual meaning.

3. G. C. D.-L. C. M. equations

The solution of a selective system is obtained as indicated in §2 from that of the dual system. Obviously, the solution of the dual system may be found by repeated applications of the solution of

$$u_1^{a_1} \cdots u_r^{a_r}\{x_1, \cdots, x_n\}_1^c = v_1^{b_1} \cdots v_s^{b_s}\{y_1, \cdots, y_m\}_2^d,$$

in which $\{\}_1$, $\{\}_2$ are unrestricted (identical or distinct, simple or compound) symbols constructed from (), [] in the same way as $\{\}'$ from ()', []' in §1; the a, b are constant integers $\geqq 0$, and c, d integers > 0; and the u, x, v, y are variable integers. Introducing new variables x_{n+1}, y_{m+1}, we first solve

$$\{x_1, \cdots, x_n\}_1 = x_{n+1}, \qquad \{y_1, \cdots, y_m\}_2 = y_{m+1},$$

as described presently. Each solution is by power products in a certain minimum number of integer parameters, subject to a set of G. C. D. conditions; say this gives

$$x_{n+1} = \phi_1^{\alpha_1} \cdots \phi_f^{\alpha_f}, \qquad y_{n+1} = \psi_1^{\beta_1} \cdots \psi_g^{\beta_g}.$$

The equation then becomes

$$u_1^{a_1} \cdots u_r^{a_r}\phi_1^{c\alpha_1} \cdots \phi_f^{c\alpha_f} = v_1^{b_1} \cdots v_s^{b_s}\psi_1^{d\beta_1} \cdots \psi_g^{d\beta_g}.$$

This is a simple multiplicative equation,[2,3] and hence its solution is obtainable non-tentatively in a finite number of steps. The solution exhibits each of the u, ϕ, v, ψ as a power product in a certain minimum number of parameters θ, subject[2] to a set of G. C. D. conditions. The G. C. D. conditions on the ϕ, ψ occurring in the previous step are replaced by their equivalents in terms of the θ, on substituting for the ϕ, ψ their expressions as power products in the θ; if any θ occurs to a power higher than the first, this power is replaced by the first power.

A system of $\{\}$-equations leads in the same way to a simple multiplicative system, and hence the solution is obtainable non-tentatively in a finite number of steps.[2]

It remains to show how an equation of the form

$$\{z_1, \cdots, z_n\} = t_{n-1}, \qquad n > 1,$$

is to be solved for $z_1, \cdots, z_n, t_{n-1}$, and we shall assume first that the symbol $\{\}$ is simple. From this we shall obtain the solution when $\{\}$ is compound by iterating the method for $\{\}$ simple.

---

[3] M. Ward, ibid., 67–76.

4. Simple G. C. D. and L. C. M. equations

The abstract identity of ()′, []′ and (), [] has been familiar for over a century, and Dedekind[4] (1897) referred to both as giving instances of his dual groups (structures, lattices), constructed on the operations now frequently called union and intersection; he also cited the classic instance, due to Boole, of logical sum and logical product. All three interpretations are required for a complete discussion of selective equations; the Boolean instance applies when we seek a minimum set of G. C. D. conditions, or their ()′-equivalents, which imply the complete set as given by the method of solution. For the moment we are concerned only with (), []; and we shall require the usual postulates

$$U(x, y) = U(y, x),$$
$$U(U(x, y), z) = U(x, U(y, z)),$$
$$U(x, C(x, y)) = x,$$

and the dual set obtained from these on replacing U by C and C by U. These are satisfied if $U(x, y)$ is the union of x, y and $C(x, y)$ their intersection; or if $U(x, y) \equiv (x, y)$, $C(x, y) \equiv [x, y]$; or if $U(x, y) \equiv (x, y)'$, $C(x, y) \equiv [x, y]'$.

The following recurrence relations are immediate consequences of the postulates,

$$(x_1, \cdots, x_n) = ((x_1, \cdots, x_{n-1}), x_n),\ n \geqq 2,$$
$$[y_1, \cdots, y_n] = [[y_1, \cdots, y_{n-1}], y_n],\ n \geqq 2;$$

and there is the well known identity $(x, y)[x, y] = xy$, which is evident from the duality in §2. It is required to solve the equations

$$(x_1, \cdots, x_n) = g_{n-1},\ n > 1;$$
$$[y_1, \cdots, y_n] = h_{n-1},\ n > 1.$$

From the recurrences and the identity we have the initial solutions ($n = 2$),

$$(x_1, x_2) = g_1:$$
$$x_1 = g_1 u_1, \quad x_2 = g_1 u_2, \quad 1 = (u_1, u_2);$$
$$[y_1, y_2] = h_1:$$
$$y_1 = k_1 v_1, \quad y_2 = k_1 v_2, \quad h_1 = k_1 v_1 v_2, \quad 1 = (v_1, v_2).$$

From the first of these and the first recurrence we find

$$(x_1, \cdots, x_n) = g_{n-1},\ n > 2:$$
$$x_1 = g_{n-1} \prod_{i=1}^{n-1} u_{2i-1}, \quad x_n = g_{n-1} u_{2n-2};$$

[4] R. Dedekind, Werke, 2, 112-114.

$$x_j = g_{n-1}u_{2j-2}\prod_{i=1}^{n-j} u_{2j+2i-3}, \qquad j = 2, \cdots, n-1;$$

$$1 = (u_1, u_2) = (u_3, u_4) = \cdots = (u_{2n-3}, u_{2n-2}),$$

in which the U's are integer parameters.

Similarly, from the second recurrence, we find

$$[y_1, \cdots, y_n] = h_{n-1}, \; n > 2:$$

$$y_1 = k_1v_1; \qquad y_t = k_{t-1}v_{2t-2}, \; t = 2, \cdots, n;$$

$$h_r = k_rv_{2r-1}v_{2r}, \; r = 1, \cdots, n-1;$$

$$1 = (v_1, v_2) = (v_3, v_4) = \cdots (v_{2n-3}, v_{2n-2}),$$

in which the integer parameters k, v are to be found from the complete solution of the simple multiplicative system[2]

$$k_sv_{2s-1}v_{2s} = k_{s+1}v_{2s+1}, \; s = 1, \cdots, n-2.$$

The typical equation has the complete solution

$$k_s = af, \qquad k_{s+1} = abc,$$
$$v_{2s-1} = bg, \qquad v_{2s+1} = fgh,$$
$$v_{2s} = ch,$$
$$1 = (f, bc) = (b, fh).$$

To solve an equation involving a compound symbol $\{\cdots, \{\}_1, \cdots, \}_n$, in which each of $\{\}_1, \cdots, \{\}_n$ is a definite one of (), [], we proceed from those $\{\}_i$ which enclose no symbol $\{\}$, say these are $\{\}_a, \cdots, \{\}_t$, and solve each of $\{\}_a = u, \cdots, \{\}_t = w$ by the above method. The $\{\}_a, \cdots, \{\}_t$ are then replaced in the original equation by the power products for $u, \cdots, w$ given by the solutions, and the process is repeated, until at the last step only a simple symbol $\{\}$ remains, when one more application of the solution for a simple equation gives the complete solution of the compound equation.

By §§2–4, the solution of the selective system in §1 is reduced to the solution of a simple multiplicative system, and this may be carried out non-tentatively in a finite number of steps.[2] After the detailed discussion, it is unnecessary to give examples (especially as those of greater interest have very long solutions), but it may be mentioned that those suggested by the Dedekind axiom, or by postulated distributivity of the operators considered, furnish interesting exercises.

An alternative method for dealing with $[y_1, \cdots, y_n]$, proposed by Lebesgue,[5] can be stated more briefly than that followed here, but actually it demands far more labor if $n > 4$. If $p_1 \equiv y_1 \cdots y_n$, and if p_j, $j > 1$, denotes the product of all the G. C. D.'s of the y's taken j at a time, Lebesgue's readily proved result is

$$[y_1, \cdots, y_n] = (p_1p_3p_5 \cdots)/(p_2p_4p_6 \cdots).$$

[5] V. A. Lebesgue, Nouv. Ann. Math., 8 (1849), 350.

A point of interest is that the G. C. D. conditions occurring in the solution of a simple multiplicative system by the method of arrays,[2] are *necessary* for the solution of a selective system. In the alternative method[3] for simple multiplicative systems, the G. C. D. conditions do not appear, and seem to have no relevance for the solution. As both methods furnish the complete solution of a simple multiplicative system, it is possible that an interpretation of selective systems in terms of the alternative method may suggest an appropriate equivalent of the G. C. D. conditions.

CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

ANNALS OF MATHEMATICS
Vol. 42, No. 4, October, 1941

SEMIGROUPS ADMITTING RELATIVE INVERSES

BY A. H. CLIFFORD

(Received December 17, 1940)

By a semigroup is meant a set S of elements $a, b, \cdots$ closed under an associative binary operation:

$$(ab)c = a(bc).$$

We shall say that S *admits relative inverses* if it satisfies the following condition.

I. *To each element a of S there exists an element e of S such that* (1) *e is an identity element of a*:

$$ea = ae = a,$$

and (2) *a possesses an inverse a' relative to e in S*:

$$aa' = a'a = e.$$

It is seen almost immediately (Theorem 1) that S admits relative inverses if and only if it is the class sum of mutually disjoint groups S_e, one to each idempotent element e of S. This fact tells us little about the structure of S since the product of two of these groups is not necessarily contained in a third, but may be scattered throughout several. (This occurs in the example given at the end of the paper, though simpler examples can be given.)

It is shown in Theorem 2, however, that S is the class sum of mutually disjoint semigroups S_α of known structure such that no such scattering takes place. Each S_α is what Rees[1] calls a completely simple semigroup without zero. The structure of such a semigroup was, in the finite case, first given by Suschkewitsch,[2] who calls it a *Kerngruppe*. Moreover it is possible to arrange these S_α in a semi-lattice[3] P such that the product $S_\alpha S_\beta$ is contained in the greatest lower bound (in P) $S_{\alpha\beta}$ of S_α and S_β. The structure of S is thus determined in the large, so to speak.

In the special case in which any two idempotent elements of S commute with each other, the simple semigroups S_α reduce to groups, coinciding with the groups S_e, and the semi-lattice P is isomorphic with the semigroup of all idempotent elements of S. In this case the structure of S is completely determined

[1] D. Rees, *On semi-groups* (Proc. Cambr. Phil. Soc. 36, 1940, 387–400).

[2] A. Suschkewitsch, *Über die endlichen Gruppen ohne das Gesetz der eindeutigen Umkehrbarkeit* (Math. Ann. 99, 1928, 30–50).

[3] A semi-lattice is a partially ordered set in which any two elements α, β have a greatest lower bound $\alpha\beta$, but not necessarily a least upper bound. F. Klein (Deutsche Math. 4, 1939, 32–43) calls it a *Halbverband*.

in Theorem 3. For every pair $\alpha > \beta$ there exists a homomorphism $\phi_{\alpha\beta}$ of the group S_α into the group S_β such that if $\alpha > \beta > \gamma$ then $\phi_{\alpha\gamma} = \phi_{\alpha\beta}\phi_{\beta\gamma}$. The product $a_\alpha b_\beta$ of any element a_α of S_α with any element b_β of S_β is then given by

$$a_\alpha b_\beta = (a_\alpha \phi_{\alpha\gamma})(b_\beta \phi_{\beta\gamma})$$

where $\gamma = \alpha\beta$. The groups S_α, the semi-lattice P, and the homomorphisms $\phi_{\alpha\beta}$ can (conversely) be chosen arbitrarily subject only to the above transitivity condition on the $\phi_{\alpha\beta}$.

In the general case, however, the situation is far more complicated. The concluding section discusses only the subsemigroup S' of S consisting of a pair S_α, S_β with $\alpha > \beta$. Theorem 4 cannot, even in this simple case, be regarded as a complete determination of S'; for it involves finding a certain mapping ϕ of S_α into the structure group of S_β the actual construction of which is not evident.

1. Decomposition into groups

LEMMA 1.1. *If a is any element of S, and e is any element of S satisfying* I, *then e is idempotent.*

PROOF: $e^2 = aa'aa' = aea' = aa' = e$.

LEMMA: 1.2. *If a is any element of S, and if e_1 and e_2 both satisfy* I, *then $e_1 = e_2$.*

PROOF: Let a_1 and a_2 be inverses of a relative to e_1 and e_2 respectively, as given by I(2):

$$aa_1 = a_1a = e_1, \qquad aa_2 = a_2a = e_2.$$

Then

$$e_1e_2 = e_1aa_2 = aa_2 = e_2,$$

$$e_1e_2 = a_1ae_2 = a_1a = e_1,$$

whence $e_1 = e_2$.

We shall say that a *belongs* to the uniquely determined idempotent element e satisfying I.

LEMMA 1.3. *The set S_e of all elements of S belonging to the idempotent element e of S is a group with identity e.*

PROOF: Let a and b be any two elements of S, and let a' and b' be inverses of a and b relative to e. Then e is evidently an identity element of ab, and $b'a'$ is an inverse of ab relative to e:

$$abb'a' = aea' = aa' = e,$$

$$b'a'ab = b'eb = b'b = e.$$

Hence $ab \in S_e$. Evidently $e \in S_e$, since $e^2 = e$, and e is an identity element of S_e. If a is any element of S_e, a' an inverse of a relative to e, and $b = ea'e$, then

$$eb = be = b,$$

$$ab = aea'e = aa'e = ee = e,$$

$$ba = ea'ea = ea'a = ee = e.$$

Hence $b \in S_e$ and b is an inverse of a therein. Thus S_e is a group.

Lemma 1.3 shows in particular that, among all possible inverses of a relative to the idempotent e to which a belongs, there is exactly one which also belongs to e. *This we shall denote by* a^{-1}.

Now by I every element of S belongs to at least one of the groups S_e, and by Lemma 1.2 to exactly one. Hence S is the class sum of the mutually disjoint groups S_e. We therefore have the following theorem,[4] the converse being evident.

THEOREM 1. *A semigroup admits relative inverses if and only if it is the class sum of mutually disjoint groups.*

2. Decomposition into completely simple semigroups

A subset $\mathfrak{a}$ of S is called an *ideal* if, for any a in $\mathfrak{a}$ and any x in S, ax and xa are in $\mathfrak{a}$; in other words, $\mathfrak{a}$ is an ideal if $SaS \subseteq \mathfrak{a}$. (We shall not have occasion to consider one-sided ideals.) The class sum $\mathfrak{a} \cup \mathfrak{b}$ and intersection $\mathfrak{a} \cap \mathfrak{b}$ of any two ideals $\mathfrak{a}$ and $\mathfrak{b}$ are evidently ideals, and likewise the set-product $\mathfrak{ab}$ consisting of all ab with a in $\mathfrak{a}$, b in $\mathfrak{b}$. Evidently

$$\mathfrak{ab} \subseteq \mathfrak{a} \cap \mathfrak{b}$$

and in particular $\mathfrak{a}^2 \subseteq \mathfrak{a}$. If a is any element of S, the set SaS of all elements xay of S is an ideal which we call the *principal ideal generated by* a. It contains a itself, since $a = eae$ by I, and hence contains aS and Sa.

LEMMA 2.1. *If an ideal* $\mathfrak{a}$ *contains a single element of the group* S_e *then it contains all of* S_e. *In particular, if* a^n *belongs to an ideal* $\mathfrak{a}$, *then* a *itself belongs to* $\mathfrak{a}$.

PROOF: Let a be an element in both $\mathfrak{a}$ and S_e. Then $\mathfrak{a}$ contains $a^{-1}a = e$. If b is any other element of S_e, then $\mathfrak{a}$ contains $be = b$. Hence $\mathfrak{a} \supseteq S_e$. The second statement then follows from the fact that a^n belongs to the same group S_e as a.

LEMMA 2.2. *If* $\mathfrak{a}$ *and* $\mathfrak{b}$ *are ideals of* S *then*

$$\mathfrak{ab} = \mathfrak{a} \cap \mathfrak{b}.$$

In particular, $\mathfrak{ab} = \mathfrak{ba}$ *and every ideal* $\mathfrak{a}$ *of* S *is idempotent:* $\mathfrak{a}^2 = \mathfrak{a}$.

PROOF: We prove the last statement first. Let a be any element of $\mathfrak{a}$. Then a^2 is in $\mathfrak{a}^2$, whence a is in $\mathfrak{a}^2$ by Lemma 2.1. Thus $\mathfrak{a} \subseteq \mathfrak{a}^2$ whence $\mathfrak{a} = \mathfrak{a}^2$. Hence if $\mathfrak{a}$ and $\mathfrak{b}$ are any two ideals of S,

$$\mathfrak{a} \cap \mathfrak{b} = (\mathfrak{a} \cap \mathfrak{b})(\mathfrak{a} \cap \mathfrak{b}) \subseteq \mathfrak{a} \cdot \mathfrak{b}$$

from which equality follows.

LEMMA 2.3. *The product of two principal ideals* SaS *and* SbS *of* S *is the principal ideal* $SabS$. *If* a *and* b *generate the same principal ideal* $\mathfrak{a}$, *then* ab *is also a generator of* $\mathfrak{a}$.

PROOF: If x is any element of S,

$$(bxa)^2 = bx \cdot ab \cdot xa \in SabS.$$

[4] This result generalizes Theorem 14 of A. R. Poole, *Finite Ova* (Amer. Jour. of Math. 59, 1937, 23–32).

By Lemma 2.1, $bxa \in SabS$. Hence

$$bSa \subseteq SabS$$

and by Lemma 2.2,

$$SaS \cdot SbS = SbS \cdot SaS \subseteq S \cdot bSa \cdot S \subseteq SabS.$$

The opposite inclusion is evident.

If $SaS = SbS = \mathfrak{a}$, then from the above and Lemma 2.2,

$$SabS = SaS \cdot SbS = \mathfrak{a}^2 = \mathfrak{a}.$$

LEMMA 2.4. *If $\mathfrak{a}$ is a principal ideal of S then the set $\mathfrak{a}'$ of all elements of $\mathfrak{a}$ which are not generators of $\mathfrak{a}$ is also an ideal of S.*

PROOF: $S\mathfrak{a}'S \subseteq SaS \subseteq \mathfrak{a}$; we wish to show that $S\mathfrak{a}'S \subseteq \mathfrak{a}'$. Suppose $S\mathfrak{a}'S$ contained an element a of $\mathfrak{a}$ not in $\mathfrak{a}'$. Then $a = xa'y$ with a' in $\mathfrak{a}'$, and a generates $\mathfrak{a}$. Hence

$$\mathfrak{a} = SaS = Sxa'yS \subseteq Sa'S;$$

thus a' would be a generator of $\mathfrak{a}$, contrary to $a' \in \mathfrak{a}'$.

An element 0 of a semigroup S is called a *zero element* of S if $0a = a0 = 0$ for all a in S. Rees (loc. cit. p. 392) defines a simple semigroup to be a semigroup S whose only ideals are S and the null ideal (0) consisting of 0 alone (if S has a zero). By a *simple semigroup without zero* we shall mean a semigroup S whose only ideal is S itself, including thereby the case in which S consists of a single element. Such a semigroup is characterized by the property that $SaS = S$ for any element a of S, or that, for given a and b in S, the equation $xay = b$ is always solvable for x and y in S.

LEMMA 2.5. *If $\mathfrak{a}$ is a principal ideal of S, then the set $S_{\mathfrak{a}}$ of all generators of $\mathfrak{a}$ is a simple semigroup without zero and admits relative inverses.*

PROOF: $S_{\mathfrak{a}}$ is a subsemigroup of S by Lemma 2.3. Let a and b be arbitrary elements of $S_{\mathfrak{a}}$; we are to show that $xay = b$ is solvable for x and y in $S_{\mathfrak{a}}$.

Now $\mathfrak{a}a\mathfrak{a}$ is an ideal containing a^3, and hence containing a by Lemma 2.1. Since a generates $\mathfrak{a}$ and $\mathfrak{a}a\mathfrak{a} \subseteq \mathfrak{a}$, we have $\mathfrak{a}a\mathfrak{a} = \mathfrak{a}$. Hence $b = xay$ with x and y in $\mathfrak{a}$. By Lemma 2.4, x and y must both lie in $S_{\mathfrak{a}}$; for if either of them lay in the ideal $\mathfrak{a}' = \mathfrak{a} - S_{\mathfrak{a}}$ then the same would be true of b, contrary to $b \in S_{\mathfrak{a}}$. Hence $S_{\mathfrak{a}}$ is a simple semigroup without zero.

If a is any element of $S_{\mathfrak{a}}$, and e is the idempotent to which a belongs, then by Lemma 2.1, $S_e \subseteq \mathfrak{a}$. By the same lemma, no element of S_e can belong to the ideal $\mathfrak{a}'$. Hence $S_e \subseteq S_{\mathfrak{a}}$. $S_{\mathfrak{a}}$ is thus the class sum of certain of the groups S_e, and so admits relative inverses.

If e and f are idempotents of S, then e is said to be *under* f (Rees p. 393) if $ef = fe = e$. We shall write $e \leqq f$ if e is under f; the relation $\leqq$ is easily seen to be a partial ordering of the set of idempotents of S. Rees makes the following definitions (p. 393).

An idempotent f is *primitive* if there exists no idempotent $e \neq 0$ such that $e < f$. A semigroup S is *completely simple* if

(i) *S is simple;*

(ii) *to each a in S there exist idempotents e and f in S such that $ea = af = a$;*

(iii)′ *every idempotent of S is primitive.* (We take here his condition (iii)′ on p. 394 which he proves to be equivalent to (iii).)

Rees (p. 399) shows that any completely simple semigroup is isomorphic with what he calls (p. 397) a *regular matrix semigroup over a group with zero.* When no zero element is present, the latter may be described as follows.

Let G be any group. Let J and K be any two sets of indices; denote the elements of J by $i, j, \cdots$, and those of K by $\kappa, \lambda, \cdots$. To each pair i, κ of indices ($i \in J, \kappa \in K$) assign an arbitrary but fixed element $p_{\kappa i}$ of G. The semigroup S shall consist of all triples ("matrices")

$$(a; i, \kappa) \qquad (a \in G; i \in J, \kappa \in K)$$

with multiplication defined by

$$(a; i, \kappa)(b; j, \lambda) = (ap_{\kappa j}b; i, \lambda). \tag{2.1}$$

The idempotents of S are readily seen from (2.1) to be the elements $(p_{\kappa i}^{-1}; i, \kappa)$.

LEMMA 2.6. *If S is a semigroup admitting relative inverses, and if e and f are idempotents of S such that $e \leqq f$ and f lies in SeS, then $e = f$.*

PROOF: From $f \in SeS$ we have $f = xey$ for some x and y in S. Setting $a = fxf$, $b = fyf$, and using $fe = ef = e$, we have

$$aeb = fxfefyf = fxeyf = fff = f$$

together with

$$fa = af = a, \qquad fb = bf = b.$$

Let g be the idempotent to which a belongs. Then

$$f = aeb = gaeb = gf = a^{-1}af = a^{-1}a = g.$$

Hence a belongs to f, and similarly b belongs to f. Consequently

$$a^{-1}fb^{-1} = a^{-1}aebb^{-1} = fef = e.$$

Since a^{-1} and b^{-1} are in S_f, this implies that e is in S_f, whence $e = f$.

LEMMA 2.7. *A simple semigroup S without zero is completely simple if and only if it admits relative inverses.*

PROOF: Assume that S admits relative inverses. Condition (i) for complete simplicity holds by hypothesis. (ii) follows from I(1), and (iii)′ from Lemma 2.6, since $SeS = S$.

Conversely, if S is completely simple then it has the matrix structure described above. If $(a; i, \kappa)$ is any element of S one readily verifies by direct calculation from (2.1) that $(p_{\kappa i}^{-1}; i, \kappa)$ is an identity and $(p_{\kappa i}^{-1}a^{-1}p_{\kappa i}^{-1}; i, \kappa)$ a relative inverse of $(a; i, \kappa)$.

By a *semi-lattice* (Klein, loc. cit.) we shall mean a commutative semigroup P

all of whose elements are idempotent. The reason for this term is that if (as above) we define $\alpha \leqq \beta$ (α and β in P) to mean $\alpha\beta = \alpha$, then P is a partially ordered set under $\leqq$ in which every pair of elements α and β have a greatest lower bound, namely $\alpha\beta$. The converse is likewise evident.

Now Lemmas 2.2 and 2.3 show that the set $\mathfrak{P}$ of principal ideals of S is a semi-lattice (actually $\mathfrak{a}\mathfrak{b} = \mathfrak{a} \cap \mathfrak{b}$ and $\mathfrak{a} \leqq \mathfrak{b}$ is equivalent to $\mathfrak{a} \subseteq \mathfrak{b}$). To each $\mathfrak{a}$ in $\mathfrak{P}$ corresponds the semigroup $S_{\mathfrak{a}}$ of all generators of $\mathfrak{a}$. If $a \in S_{\mathfrak{a}}$, $b \in S_{\mathfrak{b}}$ then by Lemma 2.3, ab generates $\mathfrak{a}\mathfrak{b}$ and so belongs to $S_{\mathfrak{a}\mathfrak{b}}$. Hence $S_{\mathfrak{a}}S_{\mathfrak{b}} \subseteq S_{\mathfrak{a}\mathfrak{b}}$. Every element a of S is a generator of exactly one principal ideal $\mathfrak{a} = SaS$ of S, and so belongs to exactly one $S_{\mathfrak{a}}$. Using Lemmas 2.5 and 2.7 we arrive at the following description of the gross structure of S.

THEOREM 2. *Every semigroup S admitting relative inverses determines a semi-lattice P such that to each element α of P there corresponds a subsemigroup S_α of S with the following properties.*

(1) *The S_α are mutually disjoint and their class sum is S.*

(2) *Each S_α is a completely simple semigroup without zero, hence a matrix semigroup over a group.*

(3) $S_\alpha S_\beta \subseteq S_{\alpha\beta}$, *where $\alpha\beta$ is the product of α and β in P.*

Conversely, any semigroup having this structure admits relative inverses.

We do not mention here the isomorphism of P with the semi-lattice of principal ideals of S. But if an ideal of S contains a single element of S_α it must by (2) contain the whole of S_α. From this it is clear that if S is any semigroup having the structure described in Theorem 2, then the lattice of all its ideals is isomorphic with the lattice of all ideals of P.

3. Case in which the idempotents of S commute

LEMMA 3.1. *If S is a semigroup admitting relative inverses, and if every pair of idempotent elements of S commute with each other, then every idempotent of S is in the center of S.*

PROOF: The set $\mathfrak{E}$ of idempotents of S evidently forms a commutative subsemigroup of S, hence a semi-lattice as defined above. We prove the lemma in four stages; e and f denote idempotents, as usual.

(1) If $SeS = SfS$ then $e = f$. For $ef \leqq e$ and $e \in SefS$ since $SefS = SeS$ by Lemma 2.3. Hence $ef = e$ by Lemma 2.6, that is, $e \leqq f$. By symmetry, $f \leqq e$, whence $e = f$.

(2) If $SaS = SeS$ then $a \in S_e$. For if $a \in S_f$ then $SaS = SfS$ by Lemma 2.1, whence $f = e$ by (1).

(3) If $a \in S_f$ and $e \leqq f$, then $ae = ea$. For $e = fe = a^{-1}ae$ and so $SeS \subseteq SaeS$. Since the opposite inclusion is evident, $SaeS = SeS$. Similarly $SeaS = SeS$. Hence ae and ea belong to S_e by (2), whence $ae = e \cdot ae = ea \cdot e = ea$.

(4) Now let e and a be arbitrary. Suppose a belongs to S_f. Since $ef \leqq f$ it follows from (3) that $a \cdot ef = ef \cdot a$. But

$$a \cdot ef = a \cdot fe = af \cdot e = ae,$$
$$ef \cdot a = e \cdot fa = ea,$$

whence $ae = ea$.

Part (2) in the above proof shows that every generator of $\mathfrak{a} = SaS$ lies in S_e, and hence $S_{\mathfrak{a}} = S_e$. Thus the simple semigroups $S_{\mathfrak{a}}$ reduce to the groups S_e in this case, and the semi-lattice $\mathfrak{P}$ is isomorphic with $\mathfrak{E}$. We shall, as in Theorem 2, denote the elements of the abstract semi-lattice P isomorphic with $\mathfrak{P}$ and $\mathfrak{E}$ by $\alpha, \beta, \gamma, \cdots$, and the elements of $\mathfrak{E}$ correspondingly by $e_\alpha, e_\beta, \cdots$. We thus have

$$e_\alpha e_\beta = e_{\alpha\beta} = e_{\beta\alpha} = e_\beta e_\alpha$$

and $e_\alpha \geqq e_\beta$ if and only if $\alpha \geqq \beta$. S_α is the same as S_{e_α} and

$$S_\alpha S_\beta \subseteq S_{\alpha\beta}.$$

In particular, if $\alpha \geqq \beta$ then $S_\alpha S_\beta \subseteq S_\beta$ and $e_\alpha e_\beta = e_\beta$. Elements of S_α will be denoted by $a_\alpha, b_\alpha, \cdots$.

LEMMA 3.2. *If $\alpha \geqq \beta$, the mapping $a_\alpha \to a_\alpha\phi_{\alpha\beta}$ defined by*

$$a_\alpha\phi_{\alpha\beta} = a_\alpha e_\beta$$

is a homomorphism of S_α into S_β. If $\alpha \geqq \beta \geqq \gamma$ then the homomorphism $\phi_{\alpha\gamma}$ of S_α into S_γ is the product of the homomorphisms $\phi_{\alpha\beta}$ and $\phi_{\beta\gamma}$ of S_α into S_β and S_β into S_γ:

$$\phi_{\alpha\gamma} = \phi_{\alpha\beta}\phi_{\beta\gamma} \qquad (\alpha \geqq \beta \geqq \gamma). \tag{3.1}$$

$\phi_{\alpha\alpha}$ is the identical automorphism of S_α.

PROOF: That $\phi_{\alpha\beta}$ maps S_α into S_β is clear. If a_α and b_α are arbitrary elements of S_α then (using Lemma 3.1)

$$(a_\alpha\phi_{\alpha\beta})(b_\alpha\phi_{\alpha\beta}) = a_\alpha e_\beta b_\alpha e_\beta = a_\alpha b_\alpha e_\beta = (a_\alpha b_\alpha)\phi_{\alpha\beta}.$$

If $\alpha \geqq \beta \geqq \gamma$ then for any a_α in S_α,

$$(a_\alpha\phi_{\alpha\beta})\phi_{\beta\gamma} = (a_\alpha e_\beta)e_\gamma = a_\alpha(e_\beta e_\gamma) = a_\alpha e_\gamma = a_\alpha\phi_{\alpha\gamma}.$$

Finally, $a_\alpha\phi_{\alpha\alpha} = a_\alpha e_\alpha = a_\alpha$, for any a_α in S_α.

LEMMA 3.3. *If a_α and b_β are any two elements of S (a_α in S_α, b_β in S_β) then*

$$a_\alpha b_\beta = (a_\alpha\phi_{\alpha\gamma})(b_\beta\phi_{\beta\gamma}) \tag{3.2}$$

where $\gamma = \alpha\beta$.

PROOF:
$$a_\alpha b_\beta = a_\alpha e_\alpha \cdot b_\beta e_\beta = a_\alpha b_\beta e_\alpha e_\beta = a_\alpha b_\beta e_\gamma$$
$$= a_\alpha e_\gamma \cdot b_\beta e_\gamma = (a_\alpha\phi_{\alpha\gamma})(b_\beta\phi_{\beta\gamma}).$$

From these two lemmas it is clear that the structure of S is completely known when we know the semi-lattice P, the groups S_α, and the homomorphisms $\phi_{\alpha\beta}$.

Conversely, let P be any semi-lattice, and to each α in P assign a group S_α such that no two of them have an element in common. Furthermore to each pair of elements α, β of P such that $\alpha > \beta$ assign a homomorphism $\phi_{\alpha\beta}$ of S_α into S_β. Define $\phi_{\alpha\alpha}$ to be the identical automorphism of S_α. Let S be the class sum of the groups S_α. Then, if the homomorphisms $\phi_{\alpha\beta}$ satisfy the transitivity conditions (3.1) and we define multiplication in S by (3.2), S is an associative semigroup.

For if a_α , b_β , c_γ are any three elements of S,

$$(a_\alpha b_\beta)c_\gamma = (a_\alpha\phi_{\alpha,\alpha\beta}\cdot b_\beta\phi_{\beta,\alpha\beta})c_\gamma$$
$$= (a_\alpha\phi_{\alpha,\alpha\beta}\cdot b_\beta\phi_{\beta,\alpha\beta})\phi_{\alpha\beta,\alpha\beta\gamma}\cdot c_\gamma\phi_{\gamma,\alpha\beta\gamma}$$

since the first factor on the right is in $S_{\alpha\beta}$. Using first the fact that $\phi_{\alpha\beta,\alpha\beta\gamma}$ is a homomorphism and then (3.1), we get

$$(a_\alpha b_\beta)c_\gamma = (a_\alpha\phi_{\alpha,\alpha\beta\gamma}\cdot b_\beta\phi_{\beta,\alpha\beta\gamma})\cdot c_\gamma\phi_{\gamma,\alpha\beta\gamma}\,.$$

Similarly,

$$a_\alpha(b_\beta c_\gamma) = a_\alpha\phi_{\alpha,\alpha\beta\gamma}\cdot(b_\beta\phi_{\beta,\alpha\beta\gamma}\cdot c_\gamma\phi_{\gamma,\alpha\beta\gamma}),$$

and these are equal by the associative law in $S_{\alpha\beta\gamma}$.

The semigroup S so constructed evidently admits relative inverses, since it is a sum of groups. Likewise it is evident that the only idempotent elements of S are the identity elements e_α of the groups S_α . Since $\phi_{\alpha\beta}$ is a homomorphism, $e_\alpha\phi_{\alpha\beta}$ must be the identity element e_β of S_β . Hence if e_α is any idempotent element of S, and b_β any element of S,

$$e_\alpha b_\beta = (e_\alpha\phi_{\alpha\gamma})(b_\beta\phi_{\beta\gamma}) = e_\gamma(b_\beta\phi_{\beta\gamma}) = b_\beta\phi_{\beta\gamma}\,,$$
$$b_\beta e_\alpha = (b_\beta\phi_{\beta\gamma})(e_\alpha\phi_{\alpha\gamma}) = (b_\beta\phi_{\beta\gamma})e_\gamma = b_\beta\phi_{\beta\gamma}\,. \qquad (\gamma = \alpha\beta)$$

Thus e_α is in the center of S.

Finally we remark that the semi-lattice P, the groups S_α , and the homomorphisms $\phi_{\alpha\beta}$ constitute a complete set of invariants of S. To see that they are invariants, suppose $S \cong \bar{S}$. Since idempotents must correspond to idempotents, $\mathfrak{E} \cong \bar{\mathfrak{E}}$, and the same abstract P may serve for both. But if $e \leftrightarrow \bar{e}$ then S_e and $\bar{S}_{\bar{e}}$ correspond, so that S_α and $\bar{S}_\alpha$ are isomorphic. Since

$$\bar{a}_\alpha\bar{\phi}_{\alpha\beta} = \bar{a}_\alpha\bar{e}_\beta \leftrightarrow a_\alpha e_\beta = a_\alpha\phi_{\alpha\beta}$$

the homomorphisms $\bar{\phi}$ are cogredient with the homomorphisms ϕ in the sense that if $a \leftrightarrow \bar{a}$ then $a\phi \leftrightarrow \bar{a}\bar{\phi}$. On the other hand, if S and $\bar{S}$ have the same invariants in the above sense, we may identify $\bar{P}$ with P and then by Lemma 3.3 the given isomorphisms between the groups S_α and $\bar{S}_\alpha$ define an isomorphism between S and $\bar{S}$.

THEOREM 3. *Every semigroup which admits relative inverses and in which every pair of idempotent elements commute is isomorphic with a semigroup S constructed as follows.*

Let P be any semi-lattice, and to each α in P assign a group S_α such that no two of them have an element in common. To each pair of elements $\alpha > \beta$ of P assign a homomorphism $\phi_{\alpha\beta}$ of S_α into S_β such that if $\alpha > \beta > \gamma$ then

$$\phi_{\alpha\beta}\phi_{\beta\gamma} = \phi_{\alpha\gamma}\,.$$

Let $\phi_{\alpha\alpha}$ be the identical automorphism of S_α. Let S be the class sum of the groups S_α, and define the product of any two elements a_α, b_β of S (a_α in S_α and b_β in S_β) by

$$a_\alpha b_\beta = (a_\alpha \phi_{\alpha\gamma})(b_\beta \phi_{\beta\gamma})$$

where $\gamma = \alpha\beta$ is the product of α and β in P.

Conversely, any semigroup S constructed in this fashion admits relative inverses, and every idempotent element of S is in the center of S.

The semi-lattice P, the groups S_α, and the homomorphisms $\phi_{\alpha\beta}$ together constitute a complete set of invariants of S.

4. Structure of a pair S_α, S_β with $\alpha > \beta$

We return now to the general case described in Theorem 2, and consider the structure of the subsemigroup S' of S consisting of a pair S_α, S_β with $\alpha > \beta$. For this purpose we may evidently assume that $S = S'$. As far as the results of this section go, S_α may be any semigroup, while S_β is to have the structure described in §2, the elements of S_β being the triples $(a; i, \kappa)$ with multiplication defined by (2.1). We shall denote the elements of S_α by capital letters $A, B, \cdots$.

Suppose that

$$A(a; i, \kappa) = (a'; i', \kappa').$$

Multiplying on the right by the idempotent $(p_{\kappa i}^{-1}; i, \kappa)$ to which $(a; i, \kappa)$ belongs, we get

$$A(a; i, \kappa) = (a' p_{\kappa' i} p_{\kappa i}^{-1}; i', \kappa),$$

whence $\kappa = \kappa'$. Hence

$$A(a; i, \kappa) = (a'; i', \kappa)$$

and similarly

$$(a; i, \kappa)A = (a''; i, \kappa'').$$

Suppose now that

$$A(p_{\kappa i}^{-1}; i, \kappa) = (a'; i', \kappa),$$
$$A(p_{\lambda i}^{-1}; i, \lambda) = (a''; i'', \lambda).$$

Multiplying the first of these on the right by $(p_{\lambda i}^{-1}; i, \lambda)$ we obtain

$$A(p_{\lambda i}^{-1}; i, \lambda) = (a' p_{\kappa i} p_{\lambda i}^{-1}; i', \lambda).$$

Hence

$$i'' = i', \qquad a'' = a' p_{\kappa i} p_{\lambda i}^{-1}.$$

The first of these shows that i' depends only on A and i and not on κ; we may

therefore write $i' = Ai$. The second shows that $a'p_{\kappa i}$ is likewise independent of κ; we may therefore write

$$a'p_{\kappa i} = A\phi_i$$

where ϕ_i is some mapping (depending on i) of S_α into the structure group G of S_β. We then have

$$A(p_{\kappa i}^{-1}\,;\, i,\, \kappa) = (A\phi_i \cdot p_{\kappa i}^{-1}\,;\, Ai\,,\, \kappa).$$

Finally, multiplying on the right by $(a;\, i,\, \kappa)$,

$$A(a;\, i,\, \kappa) = (A\phi_i \cdot a;\, Ai,\, \kappa). \tag{4.1}$$

By left-right duality, there likewise exist mappings

$$\kappa \longrightarrow \kappa A \text{ of } K \text{ into itself},$$
$$A \longrightarrow A\psi_\kappa \text{ of } S_\alpha \text{ into } G,$$

such that

$$(a;\, i,\, \kappa)A = (a \cdot A\psi_\kappa\,;\, i,\, \kappa A). \tag{4.2}$$

By double applications of (4.1) and (4.2) we find

$$AB(a;\, i,\, \kappa) = (A\phi_{Bi} \cdot B\phi_i \cdot a;\, A(Bi),\, \kappa),$$
$$(a;\, i,\, \kappa)AB = (a \cdot A\psi_\kappa \cdot B\psi_{\kappa A}\,;\, i,\, (\kappa A)B).$$

Hence

$$(AB)i = A(Bi), \qquad \kappa(AB) = (\kappa A)B; \tag{4.3}$$

$$(AB)\phi_i = (A\phi_{Bi})(B\phi_i), \qquad (AB)\psi_\kappa = (A\psi_\kappa)(B\psi_{\kappa A}). \tag{4.4}$$

Likewise

$$(a;\, j,\, \kappa)A \cdot (b;\, i,\, \lambda) = (a \cdot A\psi_\kappa \cdot p_{\kappa A, i} b;\, j,\, \lambda),$$
$$(a;\, j,\, \kappa) \cdot A(b;\, i,\, \lambda) = (ap_{\kappa, Ai} \cdot A\phi_i \cdot b;\, j,\, \lambda),$$

whence

$$(A\psi_\kappa)p_{\kappa A, i} = p_{\kappa, Ai}(A\phi_i). \tag{4.5}$$

Conversely, let S_α be any semigroup and S_β a matrix semigroup over a group G as above. Let there be given a left representation of S_α by mappings of J into itself, and a right representation of S_α by mappings of K into itself, in other words (4.3). To each i in J (and each κ in K) let there be given a mapping $\phi_i(\psi_\kappa)$ of S_α into G such that (4.4) and (4.5) hold. Let S be the class sum of S_α and S_β. Then if we define multiplication between elements of S_α and S_β by (4.1) and (4.2), the associative law holds, and S is a semigroup. Of the six possible cases to consider, three are evident from the derivation of (4.3), (4.4), (4.5), and the remaining three are identically satisfied, as is easily seen.

The equations (4.4) and (4.5) can be simplified by normalizing[5] the $p_{\kappa i}$-matrix so that

$$p_{\kappa 1} = p_{1i} = e$$

where e is the identity element of G. (No confusion arises by denoting these particular elements of J and K by the same letter 1.) Rather than refer to Rees's Theorem 2.91 and the notion of equivalence, simply observe that if we write

$$[a; i, \kappa] = (q_i a r_\kappa ; i, \kappa)$$

then

$$[a; j, \kappa][b; i, \lambda] = [a r_\kappa p_{\kappa i} q_i b; j, \lambda].$$

Thus $p_{\kappa i}$ is replaced by

$$p'_{\kappa i} = r_\kappa p_{\kappa i} q_i .$$

If we choose

$$r_\kappa = p_{11} p_{\kappa 1}^{-1} , \qquad q_i = p_{1i}^{-1}$$

then $p'_{1i} = p'_{\kappa 1} = e$.

Assuming this normalization and setting $\kappa = 1$, $i = 1$ in (4.5) we obtain $A\psi_1 = A\phi_1$. Denote $\psi_1 = \phi_1$ by ϕ. Setting $\kappa = 1$ and $i = 1$ separately in (4.5) we find

$$A\phi_i = (A\phi) p_{1A,i} , \qquad A\psi_\kappa = p_{\kappa,A1}(A\phi). \tag{4.6}$$

Putting these back in (4.5),

$$p_{\kappa,A1}(A\phi) p_{\kappa A,i} = p_{\kappa,Ai}(A\phi) p_{1A,i} . \tag{4.7}$$

Setting $i = 1$ (or $\kappa = 1$) in (4.4) and using (4.6),

$$(AB)\phi = (A\phi) p_{1A,B1}(B\phi). \tag{4.8}$$

(4.1) and (4.2) become respectively

$$A(a; i, \kappa) = (A\phi \cdot p_{1A,i} a; Ai, \kappa), \tag{4.9}$$

$$(a; i, \kappa)A = (a p_{\kappa,A1} \cdot A\phi; i, \kappa A). \tag{4.10}$$

THEOREM 4. *Let S be a semigroup which splits into two disjoint semigroups S_α and S_β such that $S_\alpha S_\beta$ and $S_\beta S_\alpha$ are contained in S_β, and where S_β is a completely simple semigroup without zero. If in the representation of S_β as a matrix semigroup over a group G we normalize the defining matrix $(p_{\kappa i})$ so that $p_{\kappa 1} = p_{1i} = e$, the identity element of G, then there exist left and right representations* (4.3) *of S_α by*

[5] That this normalization is possible was shown by Suschkewitsch (loc. cit.) in the finite case.

mappings of the index classes J and K of S_β into themselves, and a mapping ϕ of S_α into G satisfying (4.7) *and* (4.8), *such that the products in $S_\alpha S_\beta$ and $S_\beta S_\alpha$ are given by* (4.9) *and* (4.10).

Conversely, let S_α be any semigroup and S_β a completely simple semigroup without zero having no element in common with S_α. Let S_β be given as a matrix semigroup over a group G with matrix $(p_{\kappa i})$ normalized as above. Let there be given left and right representations (4.3) *of S_α by mappings of the index classes J and K of S_β into themselves, and a mapping ϕ of S_α into G satisfying* (4.7) *and* (4.8). *Let S be the class sum of S_α and S_β. Then* (4.9) *and* (4.10) *define an associative multiplication in S such that $S_\alpha S_\beta$ and $S_\beta S_\alpha$ are contained in S_β.*

PROOF: The first part has of course already been shown. To show the second part, define $A\phi_i$ and $A\psi_\kappa$ by (4.6), so that (4.9) and (4.10) go back into (4.1) and (4.2). We already have (4.3) and so need prove only (4.4) and (4.5) for associativity. Using the definition (4.6) we have

$$\begin{aligned}(AB)\phi_i &= (AB)\phi \cdot p_{1AB,i} \\ &= (A\phi)p_{1A,B1}(B\phi)p_{1AB,i}\end{aligned}$$

by (4.8). In (4.7) replace A by B and κ by $1A$:

$$p_{1A,B1}(B\phi)p_{1AB,i} = p_{1A,Bi}(B\phi)p_{1B,i}\,.$$

Hence

$$\begin{aligned}(AB)\phi_i &= (A\phi)p_{1A,Bi}(B\phi)p_{1B,i} \\ &= (A\phi_{Bi})(B\phi_i).\end{aligned}$$

The second part of (4.4) is shown in a similar manner. To show (4.5) we have by (4.7)

$$\begin{aligned}(A\psi_\kappa)p_{\kappa A,i} &= p_{\kappa,A1}(A\phi)p_{\kappa A,i}, \\ &= p_{\kappa,Ai}(A\phi)p_{1A,i} \\ &= p_{\kappa,Ai}(A\phi_i).\end{aligned}$$

That $S_\alpha S_\beta$ and $S_\beta S_\alpha$ are contained in S_β is obvious.

If all the $p_{\kappa i} = e$, which is the case if and only if the product of any two idempotents of S_β is idempotent, (4.7) is no condition at all on ϕ, and (4.8) reduces to

$$(AB)\phi = (A\phi)(B\phi).$$

Hence the structure of S is determined in this case by the representations (4.3) and a homomorphism ϕ of S_α into G.

We close with an example to show that in the contrary case ϕ need not be a homomorphism. Let G be a cyclic group of order three, $G = \{e, a, a^2\}$, $a^3 = e$. Let $J = K = \{1, 2\}$. Let $p_{11} = p_{12} = p_{21} = e$, $p_{22} = a$. Let S_α be a cyclic group of order two, $S_\alpha = \{E, A\}$, $A^2 = E$. Let

$$\begin{aligned}E1 &= 1E = 1, & E2 &= 2E = 2, \\ A1 &= 1A = 2, & A2 &= 2A = 1.\end{aligned}$$

Finally let

$$E\phi = e, \qquad A\phi = a.$$

This is evidently not a homomorphism of S_α into G. Yet the conditions (4.7) and (4.8) are easily verified. For example,

$$(AA)\phi = E\phi = e,$$

$$(A\phi)p_{1A,A1}(A\phi) = ap_{22}a = a^3 = e.$$

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

The publication of the supplementary number
is made possible by a grant from the
American Philosophical Society

ANNALS OF MATHEMATICS
Vol. 42, No. 5, December, 1941

DIE TOPOLOGIE DER LIESCHEN GRUPPEN ALS ALGEBRAISCHES PHÄNOMEN. I

VON HANS FREUDENTHAL

(Received November 15, 1940)

Daß man die Topologie der Lieschen Gruppen nicht rein gruppentheoretisch erfassen kann, ist bekannt. Desto merkwürdiger ist der zuerst von E. Cartan,[1] danach (ohne Kenntnis der Cartanschen Arbeit) andersartig von B. L. v. d. Waerden[2] bewiesene Satz:

Ein Isomorphismus zwischen zwei *halbeinfachen kompakten* Lieschen Gruppen ist notwendig stetig.

Daß die Halbeinfachheit allein die Gültigkeit des Satzes nicht gewährleistet, zeigt das laut v. d. Waerden von J. v. Neumann herrührende Beispiel:

In der Gruppe $\mathfrak{g}$ der komplexen Matrices $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ mit der Determinante 1 kann man durch unstetige Automorphismen f des Körpers der komplexen Zahlen unstetige Automorphismen

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \to \begin{pmatrix} f(\alpha) & f(\beta) \\ f(\gamma) & f(\delta) \end{pmatrix}$$

definieren.

Diese Gruppe $\mathfrak{g}$ ist nun als Liesche Gruppe von 6 reellen Parametern ($\Re_\alpha$, $\Re_\beta$, $\Re_\gamma$, $\Im_\alpha$, $\Im_\beta$, $\Im_\gamma$) einfach (bis auf einen diskreten Normalteiler), und doch fehlt dem Beispiel die rechte Überzeugungskraft. Und das kommt so: Bei der reellen Klassifikation der *reellen einfachen* nichtabelschen Gruppen[3] unterscheidet man die

erster Art, die *komplex nicht mehr einfach* (sondern direktes Produkt zweier konjugierter einfacher Gruppen sind), und die

zweiter Art, die *auch komplex einfach* sind.

Die erster Art lassen sich, wie man leicht beweist,[4] alle so erzeugen: Man nehme eine reelle einfache nichtabelsche Gruppe, setze sie ins Komplexe fort und trenne Real- und Imaginärteil; diesen Prozeß, der die Dimensionszahl verdoppelt, wollen wir "Verdoppelung" nennen.

Eine Gruppe erster Art haben wir gerade in obigem Beispiel angetroffen. So

[1] E. Cartan, Sur les représentations linéaires des groupes clos [Comment. Helvet. **2** (1930), 269–283].

[2] B. L. van der Waerden, Stetigkeitssätze für halbeinfache Liesche Gruppen [Math. Zeitschr. **36** (1933), 780–786].

[3] E. Cartan, Les groupes réels simples, finis et continus [Ann. Ecole Norm. Sup. (3) **31** (1914), 263–355].

[4] a. a. O.,[3] 266–267.

beschleicht uns der Verdacht, daß für den Cartan-v.d.Waerdenschen Satz nicht die Kompaktheit, sondern die Zugehörigkeit der einfachen Gruppe zur zweiten Art verantwortlich sei.

In der Tat beweisen wir den

HAUPTSATZ I: *Ein Isomorphismus zwischen zwei Lieschen Gruppen, von denen die eine einfach und von zweiter Art ist, ist notwendig stetig. Oder: Im Bereiche der Lieschen Gruppen ist die Topologie der einfachen Gruppen zweiter Art ein rein algebraisches Phänomen.*

(In der letzten Formulierung ist der Hauptsatz auch sinnvoll für jemanden, der unstetige Abbildungen überhaupt ablehnt—etwa aus intuitionistischen Erwägungen.)

BEMERKUNG 1: Der Hauptsatz bleibt gültig, wenn man die Einfachheit nur im Kleinen fordert.

BEMERKUNG 2: Hauptsatz I läßt sich in naheliegender Weise auf halbeinfache Gruppen ausdehnen (siehe v. d. Waerden, a.a.O.,[2] S. 785).

BEMERKUNG 3: Für die reellen projektiven Gruppen ist Hauptsatz I bereits von O. Schreier und B. L. van der Waerden bewiesen worden.[4a]

Man wird vermuten, daß die Aussage von Hauptsatz I für keine einfache Gruppe erster Art richtig sei; auf die nicht so ganz leichte Aufzählung aller Isomorphismen für Gruppen erster Art werde ich bald zurückkommen. Ich beabsichtige, dieselben Fragen auch für nicht-halbeinfache Gruppen zu beantworten; für die Gültigkeit von Hauptsatz I ist nämlich etwas wie die Zugehörigkeit zur zweiten Art wichtiger als die Halbeinfachheit.

Den Gedankengang unseres Beweises erkennt man leicht aus der Gliederung der Arbeit in die Sätze 1 – 9. Nur auf Satz 1 und seine wohl an und für sich interessanten Anwendungen will ich besonders hinweisen.

ALLGEMEINE BEZEICHNUNGEN

Gruppen: kleine gotische Buchstaben.
Elemente von Gruppen: kleine lateinische Buchstaben.
Gruppen-Eins: 1.
Infinitesimalgruppen: große gotische Buchstaben.
Elemente von Infinitesimalgruppen: große lateinische Buchstaben.
Mengen von Gruppen: kleine fette gotische Buchstaben.
Mengen von Infinitesimalgruppen: große fette gotische Buchstaben.
Meistens werden Gruppe und zugehörige Infinitesimalgruppe mit demselben Buchstaben bezeichnet (klein und groß).
$\cap$ = Durchschnitt, $\cup$ = Vereinigung.

Vorbereitungen—Liesche Gruppen

2. Liesche Gruppe ist bei uns jede Gruppe, die von einer Infinitesimalgruppe erzeugt wird. Wir finden es zweckmäßig, von einer Lieschen Untergruppe *nicht* (wie es manchmal geschieht) Abgeschlossenheit zu verlangen.

Unter einer *Basis* einer Infinitesimalgruppe $\mathfrak{G}$ verstehen wir ein System, aus

[4a] Die Automorphismen der projektiven Gruppen [Abh. Hamburg **6**(1928), 303-322].

dem sich alle Elemente von $\mathfrak{G}$ eindeutig linear kombinieren lassen; ein *Erzeugendensystem* von $\mathfrak{G}$ sei dagegen ein System, aus dem sich alle Elemente von $\mathfrak{G}$ durch lineare Kombination und Kommutatorbildung herstellen lassen.

3. In Lieschen Gruppenkeimen verwendet man öfters geodätische Koordinatensysteme: Ist $X_1, \cdots, X_r$ eine Basis von $\mathfrak{G}$, so erhält der Punkt

$$\exp\left(\sum_\nu \lambda^\nu X_\nu\right)$$

die Koordinaten

$$\lambda^1, \cdots, \lambda^r.$$

Alle geodätische Koordinatensysteme transformieren sich ineinander analytisch mit nicht verschwindender Funktionaldeterminante.

Die Koordinatentransformation, bei der der Punkt

$$\exp\left(\mu^1 X_1 + \cdots + \mu^{\nu_1} X_{\nu_1}\right) \cdot \exp\left(\mu^{\nu_1+1} X_{\nu_1+1} + \cdots + \mu^{\nu_2} X_{\nu_2}\right) \cdots$$
$$\cdots \exp\left(\mu^{\nu_p+1} X_{\nu_p+1} + \cdots + \mu^r X_r\right)$$

die Koordinaten

$$\mu^1, \cdots, \mu^r$$

erhält, ist eine analytische Abbildung des λ-Raums in den μ-Raum mit der Funktionaldeterminante 1 im Nullpunkt.

4. In der Lieschen Gruppe $\mathfrak{g}$ seien $a_\nu(t)$ $(\nu = 1, \cdots, r)$ r stetig differenzierbare Bögen mit $a_\nu(0) = 1$; die Tangentialvektoren $X_\nu = a_\nu'(0)$ mögen eine Basis von $\mathfrak{G}$ bilden. Dann überdecken die

$$a_1(t_1) \cdots a_r(t_r)$$

eine Umgebung von 1.

Denn die Abbildung des $(t_1, \cdots, t_r)$-Raumes in den λ-Raum, die durch

$$a_1(t_1) \cdots a_r(t_r) = \exp\left(\lambda^1 X_1 + \cdots + \lambda^r X_r\right)$$

definiert ist, besitzt in 1 die Funktionaldeterminante 1.

5. Reguläre Schichtung einer r-dim. analytischen Mannigfaltigkeit in der Umgebung eines ihrer Punkte nennen wir das analytische Bild (Funktionaldeterminante $\neq 0$) der Schichtung

$$(*) \qquad \xi_{q+1} = \text{const}, \cdots, \xi_r = \text{const}$$

des $(\xi_1, \cdots, \xi_r)$-Raumes; die Bilder der q-dim. Hyperebenen (*) heißen die Schichten.

Wir zeigen: Ein stetig differenzierbarer Bogen, der in jedem seiner Punkte a die durch a laufende Schicht einer regulären Schichtung berührt, verläuft ganz in *einer* Schicht. Es genügt, das zu beweisen für die Schichtung (*) des

ξ-Raumes. In diesem Falle gilt aber, wenn $\xi_\nu(t)$ $(\nu = 1, \cdots, r)$ die r Koordinaten des Kurvenpunktes sind, $\xi'_\nu(t) = 0$ für alle $\nu > q$, also in der Tat $\xi_\nu(t) =$ const für alle $\nu > q$, w.z.b.w.

6. Die (Rechts-oder Links-)Nebengruppenschichtung eines Lieschen Gruppenkeimes $\mathfrak{g}$ in bezug auf einen Lieschen Untergruppenkeim $\mathfrak{h}$ ist in der Umgebung jedes Punktes regulär. Anwendung von 5 und des Borelschen Überdeckungssatzes liefert: Eine stetig differenzierbare Kurve in $\mathfrak{g}$, die in jedem ihrer Punkte a die Nebengruppe $a\mathfrak{h}$ berührt, verläuft ganz in *einer* Nebengruppe.

7. In der Lieschen Gruppe $\mathfrak{g}$ seien $c_\nu(t)$ $(\nu = 1, \cdots, n)$ analytische Bögen, $c_\nu(0) = 1$. Die kleinste Gruppe, die diese Bögen enthält, heiße $\mathfrak{h}(\subset \mathfrak{g})$. $\mathfrak{h}$ ist notwendig Liesch.

Beweis: $a(t) = t^p X + t^{p+1} \cdot$(Potenzreihe in t) sei die MacLaurin-Entwicklung eines beliebigen analytischen Bogens aus $\mathfrak{h}$ mit $a(0) = 1$ und $X \neq 0$; der Vektor X heiße der "Hauptteil" von $a(t)$. $\mathfrak{H}$ sei die Menge aller solcher Hauptteile mit $a(t) < \mathfrak{h}$ nebst dem Nullvektor. $\mathfrak{H}$ ist eine Infinitesimalgruppe, denn sind X bzw. Y die Hauptteile von $a(t)$ bzw. $b(t)$, so sind $\alpha X + \beta Y$ bzw. $[X, Y]$ die Hauptteile der Bögen $a(\alpha t) \cdot b(\beta t)$ bzw. $a(t)b(t)a^{-1}(t)b^{-1}(t)$.

Sei a ein willkürliches Element von $\mathfrak{h}$; dann ist

$$a = c_{\nu_1}(t_1)^{\pm 1} \cdots c_{\nu_q}(t_q)^{\pm 1},$$

wo $\nu_1, \cdots, \nu_q$ gewisse der Zahlen $1, \cdots, n$ (evtl. mit Wiederholungen) sind. Der analytische Bogen

$$a(t) = c_{\nu_1}(tt_1)^{\pm 1} \cdots c_{\nu_q}(tt_q)^{\pm 1} \qquad (0 \leqq t \leqq 1)$$

verbindet 1 mit a in $\mathfrak{h}$. Für festes t ist

$$b(\tau) = a(t)^{-1} a(\tau + t) \qquad (-t \leqq \tau \leqq 1 - t)$$

ein analytischer Bogen in $\mathfrak{h}$; nach der Definition von $\mathfrak{H}$ gehört sein Tangentialvektor für $\tau = 0$ zu $\mathfrak{H}$. Also gehört $a'(t)$ (das ist nämlich der Tangentialvektor von $a(\tau + t)$ für $\tau = 0$) zu $a(t)\mathfrak{H}$. Sei $\mathfrak{h}_0$ die von $\mathfrak{H}$ erzeugte (Liesche) Gruppe; nun berührt $a(t)$ in jedem seiner Punkte die Nebengruppe $a(t)\mathfrak{h}_0$, liegt also nach 6 ganz in *einer* Nebengruppe, und da $a(0) = 1$ ist, ganz in $\mathfrak{h}_0$. Also ist $a = a(1) \in \mathfrak{h}_0$. Da a beliebig in $\mathfrak{h}$ war, haben wir

$$\mathfrak{h} \subset \mathfrak{h}_0. \tag{1}$$

Sei nun $X_1, \cdots, X_r$ eine Basis von $\mathfrak{H}$; definitionsgemäß existieren in $\mathfrak{h}$ analytische Bögen $a_\nu(t)$ mit den Hauptteilen X_ν:

$$a_\nu(t) = t^{p_\nu} X_\nu + \cdots.$$

In $a_\nu(t)$ führen wir als neuen Parameter $s = t^{p_\nu}$ ein,

$$a_\nu(t) = b_\nu(s);$$

$b_\nu(s)$ ist stetig differenzierbar und sein Tangentialvektor für $s = 0$ ist X_ν. Die Menge der

$$b_1(s_1) \cdots b_r(s_r)$$

ist in $\mathfrak{h}$ enthalten und überdeckt nach 4 eine Umgebung von 1 in $\mathfrak{h}_0$. Also enthält $\mathfrak{h}$ eine Umgebung von 1 in $\mathfrak{h}_0$, also ganz $\mathfrak{h}_0$:

$$\mathfrak{h} \supset \mathfrak{h}_0. \tag{2}$$

Also ((1) und (2)) ist $\mathfrak{h} = \mathfrak{h}_0$ Liesch, w.z.b.w.

8. SATZ 1: *In der Untergruppe* $\mathfrak{e}$ *der Lieschen Gruppe* $\mathfrak{g}$ *sei jedes Element analytisch verbindbar mit* 1. *Dann ist* $\mathfrak{e}$ *Liesch.*

BEWEIS: $\mathfrak{f}$ sei Liesche Untergruppe maximaler Dimension von $\mathfrak{e}$; $X_1, \cdots, X_p$ sei eine Basis der Infinitesimalgruppe von $\mathfrak{f}$. Sei a beliebig in $\mathfrak{e}$; nach Voraussetzung existiert ein analytischer Bogen $a(t)$ $(0 \leqq t \leqq 1)$ mit $a(0) = 1$, $a(1) = a$. Nach 7 erzeugen $a(t)$ und $\exp(tX_\nu)$ $(\nu = 1, \cdots, p)$ zusammen eine Liesche Untergruppe von $\mathfrak{g}$, die wegen der Maximalität (s.o.) gleich $\mathfrak{f}$ ist. Also ist $a \in \mathfrak{f}$, also $\mathfrak{e} \subset \mathfrak{f}$, also $\mathfrak{e} = \mathfrak{f}$ Liesch, w.z.b.w.

9. ANWENDUNGEN VON SATZ 1:

1. Sei $\mathfrak{k}$ Teilmenge der Lieschen Gruppe $\mathfrak{g}$; sei $\mathfrak{h}$ eine Teilmenge von $\mathfrak{g}$, die 1 enthält, und in der jedes Element mit 1 analytisch verbindbar sei (z.B. eine Liesche Untergruppe von $\mathfrak{g}$). Die von den Kommutatoren $k^{-1}h^{-1}kh$ (mit $k \in \mathfrak{k}$ und $h \in \mathfrak{h}$) erzeugte Gruppe $\mathfrak{e}$ ist Liesch.—Denn ein Kommutator $k^{-1}h^{-1}kh$ ist mit 1 durch $k^{-1}h(t)^{-1}kh(t)^{-1}$ analytisch verbindbar, wenn $h(t)$ eine analytische Verbindung von 1 mit h ist; ist $e \in \mathfrak{e}$, so ist $e = e_1 \cdots e_q$, wo die e_ν von der Gestalt $k^{-1}h^{-1}kh$ sind; ist e_ν durch $e_\nu(t)$ mit 1 verbindbar, so auch e durch $e_1(t) \cdots e_q(t)$. Die Voraussetzungen von Satz I sind also erfüllt.

2. Ein Normalteiler $\mathfrak{e}$ der Lieschen Gruppe $\mathfrak{g}$, der mit seiner Kommutatorgruppe identisch ist, ist Liesch.—Man setze nämlich in 9.1: $\mathfrak{h} = \mathfrak{g}$, $\mathfrak{k} = \mathfrak{e}$.

3. Ein Normalteiler $\mathfrak{e}$ der Lieschen Gruppe $\mathfrak{g}$, der nicht ganz im Zentrum von $\mathfrak{g}$ liegt, enthält eine Liesche Untergruppe positiver Dimension.—Ist $a \in \mathfrak{e}$, aber a nicht im Zentrum von $\mathfrak{g}$, so ist die von den $g^{-1}a^{-1}ga$ $(g \in \mathfrak{g})$ erzeugte Gruppe von positiver Dimension und nach Satz I (auch bereits nach 7) Liesch.

4. Ist die Liesche Gruppe $\mathfrak{g}$ einfach im Kleinen, so liegen alle ihre Normalteiler $(\neq \mathfrak{g})$ in ihrem (notwendig diskreten) Zentrum.[5]—Wäre nämlich $\mathfrak{e}$ ein Normalteiler, der nicht ganz im Zentrum läge, so wäre, wie aus Hauptsatz I leicht folgt, die maximale Liesche Untergruppe $\mathfrak{f}$ von $\mathfrak{e}$ ebenfalls Normalteiler von $\mathfrak{g}$ und nach 9.3 von positiver Dimension, was der Einfachheit im Kleinen widerspricht.

5. Man kann 9.4 auf im Kleinen halbeinfache Gruppen ausdehnen.

---

[5] Verallgemeinerung eines Satzes von v. d. Waerden, a. a. O.,[2] 784, für kompakte Gruppen.

10. Wir erwähnen für spätere Anwendungen den Satz: Bei einer abgeschlossenen Untergruppe einer Lieschen Gruppe ist die Menge der Komponenten (abzählbar) diskret und die Komponente von 1 eine Liesche Gruppe.[6]

Vorbereitungen—halbeinfache Gruppen[7]

11. 1. *Rang* l von $\mathfrak{G}$ nennt man bekanntlich die Zahl der algebraisch unabhängigen Wurzeln der Elemente von $\mathfrak{G}$. Bei einer halbeinfachen Gruppe ist der Rang gleich der Zahl der identisch verschwindenden Wurzeln. *Regulär* heißt ein Element von $\mathfrak{G}$, das nicht mehr als l verschwindende Wurzeln besitzt. *Reguläre* Untergruppe der halbeinfachen Gruppe $\mathfrak{G}$ heißt die Gruppe $\mathfrak{H}$ aller mit einem regulären Element vertauschbaren Elemente; auch die von einem regulären $\mathfrak{H}$ erzeugte endliche Gruppe $\mathfrak{h}$ wird regulär genannt.

In Bezug auf eine reguläre Untergruppe $\mathfrak{H}$ (Elemente H) kann man eine halbeinfache Gruppe bekanntlich auf folgende einfache Gestalt bringen:

$$[H, E_\alpha] = \alpha E_\alpha ,$$
$$[E_\alpha , E_{-\alpha}] = H_\alpha ,$$
$$[E_\alpha , E_\beta] = N_{\alpha,\beta} E_{\alpha+\beta} (\text{falls } \alpha + \beta \neq 0).$$

Die $\alpha, \beta, \cdots$ sind dabei lineare Funktionale von $\mathfrak{H}$, die "Wurzelformen"[8] von $\mathfrak{G}$ in bezug auf $\mathfrak{H}$ (die man auch als kontravariante Vektoren im Raume der H auffassen kann). Mit α ist auch $-\alpha$ Wurzelform; unter den Wurzelformen gibt es l linear unabhängige.

2. Wir setzen im Folgenden $\mathfrak{G}$ und $\mathfrak{H}$ stets als reell voraus, gehen aber, wo es nötig ist, ohneweiteres ins Komplexe über. Dann tritt mit α auch $\bar{\alpha}$ (das konjugiert komplexe Funktional) als Wurzelform auf; wir wählen E_α und $E_{\bar{\alpha}}$ so, daß sie konjugiert komplex sind.

Wegen der Halbeinfachheit ist die quadratische Form

$$\varphi(T, T) = \text{Spur der Abbildung } X \to [T[TX]]$$
$$= \sum \rho^2 + \sum N_\rho \tau^\rho \tau^{-\rho} \qquad (T = H + \sum \tau^\rho E_\rho)$$

(Summen über alle Wurzelformen erstrecken!) bekanntlich nicht entartet. N_ρ = Spur der Abbildung $X \to [E_{-\rho}[E_\rho X]]$, also $N_{\bar{\rho}} = \overline{N_\rho}$. Die Normierung[9]

---

[6] E. Cartan, La théorie des groupes finis et continus et l'analysis situs [Mémorial sc. math. XLII], p. 24.

[7] Siehe hierzu: E. Cartan, Sur la structure des groupes de transformations finis et continus [Thèse, Paris 1894].—H. Weyl, Theorie der Darstellungen kontinuierlicher halbeinfacher Gruppen durch lineare Transformationen [Math. Zeitschr. **23**(1925), 271-309; **24** (1926), 328-395], insb. Kap. III.—Wir richten uns hier nach B. L. v. d. Waerden, Die Klassifikation der einfachen Lieschen Gruppen [Math. Zeitschr. **37**(1933), 446-462].

[8] Im Interesse der Deutlichkeit unterscheiden wir zwischen den Wurzeln—das sind die charakteristischen Wurzeln der Matrices der adjungierten Gruppe, entweder für ein einzelnes Element oder als Funktionen auf der ganzen Gruppe—und den Wurzelformen—das sind die Wurzeln, als lineare Funktionen auf $\mathfrak{H}$ definiert.

[9] Siehe v. d. Waerden, a. a. O.,[7] 447.

$N_\rho = -1$ läßt sich daher so durchführen, daß E_ρ und $E_{\bar\rho}$ konjugiert komplex bleiben.

Diese Normierung hat zur Folge, daß H_ρ linear von ρ abhängt. Ist

$$\lambda = \sum_\rho p_\rho \rho$$

eine lineare Kombination von Wurzelformen, so können wir dann widerspruchslos definieren

$$H_\lambda = \sum_\rho p_\rho H_\rho ;$$

alle Elemente von $\mathfrak{H}$ lassen sich auf diese (für uns recht bequeme) Weise darstellen.

Wir haben nun

$$[H_\lambda E_\rho] = -(\rho, \lambda) E_\rho ;$$

$-(\rho, \lambda)$ ist der Wert des Funktionals ρ für den Vektor H_λ. Hier ist (λ, μ) nichts Anderes als die zu der quadratischen Form $Q(H_\lambda, H_\lambda) = \sum \rho^2$ gehörige symmetrische Bilinearform.

Da mit ρ auch $\bar\rho$ Wurzelform ist, ist die Bilinearform (λ, μ) reell (d.h. reell für reelle Argumente). Die quadratische Form (λ, λ) ist im Allgemeinen nicht definit; trotzdem werden wir (mit der nötigen Vorsicht) Bezeichnungen wie orthogonal (in Bezug auf diese quadratische Form) gebrauchen; ein Vektor $H_\lambda \neq 0$ mit $(\lambda, \lambda) = 0$ heißt isotrop; die H_ρ, für die ρ eine Wurzelform ist, können nie isotrop sein, für Wurzelformen ρ gilt: $(\rho, \rho) < 0$.

3. Mit α und β ist auch $\alpha - 2\dfrac{(\alpha, \beta)}{(\beta, \beta)}\beta$ Wurzelform; die $2\dfrac{(\alpha, \beta)}{(\beta, \beta)}$ sind ganz. Mit α, β und $\alpha + k\beta$ ist auch $\alpha + k'\beta$ Wurzelform für $0 \leqq k' \leqq k$.

12. Eine halbeinfache Gruppe ist bekanntlich dann und nur dann komplex einfach, wenn sich je zwei Wurzelformen α, β durch eine "Kette"

$$\alpha = \alpha_1 , \qquad \alpha_2 , \cdots , \alpha_p , \qquad \alpha_{p+1} = \beta$$

von Wurzelformen verbinden lassen,

$$(\alpha_\nu , \alpha_{\nu+1}) \neq 0.$$

Wir behaupten:

Ist $\mathfrak{G}$ halbeinfach und komplex einfach, und sind α, β Wurzelformen mit $(\alpha, \beta) \neq 0$, so existiert eine Wurzelform γ mit

$$(\alpha, \gamma) \neq 0, \qquad (\beta, \gamma) \neq 0.$$

Zum Beweise nehmen wir an, obige Kette sei minimal; wir müssen dann zeigen, daß $p = 2$ ist. Wir dürfen weiter annehmen

$$\frac{(\alpha_\nu, \alpha_{\nu+1})}{(\alpha_{\nu+1}, \alpha_{\nu+1})} < 0 \quad \text{für alle } \nu$$

(andernfalls ersetzen wir gewisse Wurzelformen durch ihre negativen). Aus der Minimalität folgt

$$(\alpha_\mu , \alpha_\nu) = 0 \text{ für } |\mu - \nu| > 1.$$

Sei bereits für ein $q < p$ bewiesen, daß $\sum_{\nu=2}^{q} \alpha_\nu$ entweder Wurzelform oder 0 ist; wir beweisen dasselbe für $\sum_{\nu=2}^{q+1} \alpha_\nu$. In der Tat ist

$$\frac{\left(\sum_{\nu=2}^{q} \alpha_\nu , \alpha_{q+1}\right)}{(\alpha_{q+1} , \alpha_{q+1})} = \frac{(\alpha_q , \alpha_{q+1})}{(\alpha_{q+1} , \alpha_{q+1})} < 0;$$

ist $\sum_{\nu=2}^{q} \alpha_\nu$ Wurzelform und $\neq -\alpha_{q+1}$, so ist nach 11.3 $\sum_{\nu=2}^{q} \alpha_\nu + \alpha_{q+1}$ Wurzelform; ist $\sum_{\nu=2}^{q} \alpha_\nu = -\alpha_{q+1}$, so ist $\sum_{\nu=2}^{q} \alpha_\nu + \alpha_{q+1} = 0$; ist schließlich $\sum_{\nu=2}^{q} \alpha_\nu = 0$, so ist $\sum_{\nu=2}^{q} \alpha_\nu + \alpha_{q+1} = \alpha_{q+1}$, also Wurzelform. Damit ist die Induktionsbehauptung bewiesen.

Aus ihr folgt: $\gamma = \sum_{\nu=2}^{p} \alpha_\nu$ ist Wurzelform oder 0.

$$(\alpha, \gamma) = (\alpha_1 , \gamma) = (\alpha_1 , \alpha_2) \neq 0, \qquad (\beta, \gamma) = (\alpha_{p+1} , \gamma) = (\alpha_{p+1} , \alpha_p) \neq 0.$$

γ ist demnach die gesuchte Wurzelform.

13. Aus dem Vorigen folgt: Ist $\mathfrak{G}$ halbeinfach und komplex einfach, so sind alle Zahlen (α, β) (α, β Wurzelformen) rationale Vielfache einer unter ihnen.

14. $\mathfrak{H}$ sei eine p-dim. abelsche Untergruppe der halbeinfachen Gruppe $\mathfrak{G}$; ferner sei $\mathfrak{H}$ von keiner größeren Untergruppe von $\mathfrak{G}$ Normalteiler. Dann ist $\mathfrak{H}$ regulär.

BEWEIS: Nach Voraussetzung gilt für jedes $X \in \mathfrak{G}$ mit

$$[HX] \in \mathfrak{H} \text{ für alle } H \in \mathfrak{H}$$

notwendig

$$X \in \mathfrak{H}.$$

Wir können daher $\mathfrak{G}$ auf die Gestalt

$$\begin{aligned} [HE_\alpha] &= \alpha E_\alpha \\ [HE'_\alpha] &= \alpha E'_\alpha \bmod E_\alpha \\ &\vdots \\ [E_\alpha E_{-\alpha}] &\in \mathfrak{H} \\ &\vdots \end{aligned}$$

bringen, wo die α die nichtverschwindenden Wurzelformen sind. Aus der Tatsache, daß die Spur von

$$X \to [T[TX]]$$

eine nichtentartete quadratische Form ist, schließt man wie üblich, daß $\sum \rho^2$ nicht ausgeartet ist, und daß es genau p linear unabhängige Wurzelformen gibt, also $p \leqq l$ (denn die Wurzeln sind Funktionen auf ganz $\mathfrak{G}$, die Wurzelformen nur auf $\mathfrak{H}$, und für die (linearen) Wurzelformen fallen lineare und algebraische Abhängigkeit zusammen; siehe auch 11.1). Andererseits ist p definitionsgemäß auch die Zahl der verschwindenden Wurzelformen, also $p \geqq l$. Demnach $p = l$; jedes genügend allgemeine Element von $\mathfrak{H}$ ist also regulär, und da $\mathfrak{H}$ sicher maximal abelsch ist, ist auch $\mathfrak{H}$ regulär.

15. Eine Teilmenge der Wurzelformen der halbeinfachen Gruppe nennen wir geschlossen, wenn sie mit α auch $-\alpha$ enthält und mit α und β auch $\alpha + \beta$, falls das eine Wurzelform ist.

Sei $\mathbf{P}$ eine geschlossene Menge von Wurzelformen und

$$\sigma = \sum_{\rho} p_\rho \rho$$

eine lineare Kombination aus $\mathbf{P}$, die auf ganz $\mathbf{P}$ orthogonal sei. Dann ist $\sigma = 0$.

Beweis: Die obige Darstellung von σ sei so gewählt, daß $\sum | p_\rho |$ minimal positiv ist. Es gebe ein p_α mit $p_\alpha \neq 0$.

$$(\alpha, \sigma) = \sum p_\rho (\alpha, \rho).$$

Wegen $(\alpha, \sigma) = 0$ und $p_\alpha(\alpha, \alpha) \neq 0$ gibt es ein $\beta (\neq -\alpha)$ mit

$$\operatorname{sgn} p_\alpha(\alpha, \alpha) = -\operatorname{sgn} p_\beta(\alpha, \beta),$$

$$-\operatorname{sgn} \frac{(\alpha, \beta)}{(\alpha, \alpha)} = \operatorname{sgn} \frac{p_\alpha}{p_\beta}.$$

Sei $\operatorname{sgn} p_\alpha = \operatorname{sgn} p_\beta$. Dann ist $\operatorname{sgn} \dfrac{(\alpha, \beta)}{(\alpha, \alpha)} < 0$, also nach 11.3 $\beta + \alpha$ Wurzelform $\epsilon \mathbf{P}$.

$$\sigma = \sum p_\rho \rho \pm \alpha \pm \beta \mp (\alpha + \beta) = \sum q_\rho \rho;$$

$$q_\alpha = p_\alpha \pm 1,\ q_\beta = p_\beta \pm 1,\ q_{\alpha+\beta} = p_{\alpha+\beta} \mp 1$$

(untere Zeichen für $p_\alpha > 0$, obere für $p_\alpha < 0$). Das ist eine neue Darstellung für σ, und es ist $\sum | q_\rho | < \sum | p_\rho |$. Das ist ein Widerspruch zur Minimalität.

Analog schließt man bei $\operatorname{sgn} p_\alpha = -\operatorname{sgn} p_\beta$.

16. Sei $\mathbf{P}$ eine geschlossene Menge von Wurzelformen und S das System aller linearen Relationen in $\mathbf{P}$. Dann bilden die Relationen der Form

(*) $$\alpha + \beta = 0 \quad \text{und} \quad \alpha + \beta + \gamma = 0$$

aus S ein Erzeugendensystem für S. ($\alpha, \beta, \gamma \epsilon \mathbf{P}$.)

BEWEIS: Das von (*) erzeugte Relationensystem heiße S^*. Ist $S^* \neq S$, so wählen wir die Relation

$$\sum p_\rho \rho = 0$$

in S, aber nicht in S^*, so daß $\sum | p_\rho |$ minimal ist. Wir schließen dann wörtlich wie oben auf die Existenz einer Relation

$$\sum q_\rho \rho = 0$$

in S, aber nicht in S^*, mit

$$\sum | q_\rho | < \sum | p_\rho |,$$

also auf einen Widerspruch.

17. Die Untergruppe $\mathfrak{E}$ der halbeinfachen Gruppe $\mathfrak{G}$ sei von einer Menge von Paaren E_α, $E_{-\alpha}$ aus $\mathfrak{G}$ erzeugt. Dann ist $\mathfrak{E}$ halbeinfach und $\mathfrak{E} \cap \mathfrak{H}$ reguläre Untergruppe von $\mathfrak{E}$.

BEWEIS: Ein Erzeugendensystem von $\mathfrak{E}$ besteht aus den E_α, $E_{-\alpha}$, H_α, wo α eine geschlossene Menge **A** durchläuft. Man berechne $\varphi(T, T)$ für die Gruppe $\mathfrak{E}$. Da N_α ein positives Vielfaches von $(\alpha, \alpha) \neq 0$ ist (siehe Weyl, a.a.O., S.366), brauchen wir uns nur um $\sum \rho^2$ zu kümmern, um zu beweisen, daß $\varphi(T, T)$ nicht ausgeartet ist. Für $\sum \rho^2$ folgt das aber nun aus 15, da danach kein von **A** abhängiger Vektor auf ganz **A** senkrecht steht.

Rein gruppentheoretische Definition der regulären Untergruppen

Im Großen werden wir die halbeinfachen Gruppen jetzt immer als zentrumfrei voraussetzen.

18. $\mathfrak{h}$ sei reguläre Untergruppe der halbeinfachen Gruppe $\mathfrak{g}$, $\mathfrak{h}^*$ sei die maximalabelsche Untergruppe von $\mathfrak{g}$ die $\mathfrak{h}$ enthält. Dann ist $\mathfrak{h}^*/\mathfrak{h}$ endlich.[10]

BEWEIS: Sei $a \in \mathfrak{h}^*$. Der Automorphismus

$$A \cdots x \to a^{-1}xa$$

läßt $\mathfrak{h}$ elementweise fest. A ist auch ein (reeller) Automorphismus für die komplexe Infinitesimalgruppe $\mathfrak{G}$ von $\mathfrak{g}$:

$$[H, E_\alpha] = \alpha E_\alpha, \qquad [H, AE_\alpha] = \alpha AE_\alpha, \qquad \text{also } AE_\alpha = \nu_\alpha E_\alpha.$$

(1) $$[E_\alpha, E_{-\alpha}] = H_\alpha, \qquad [AE_\alpha, AE_{-\alpha}] = H_\alpha, \qquad \text{also } \nu_\alpha \nu_{-\alpha} = 1.$$

(2) $$[E_\alpha, E_\beta] = N_{\alpha,\beta} E_{\alpha+\beta}, \; [AE_\alpha, AE_\beta] = N_{\alpha,\beta} AE_{\alpha+\beta}, \text{ also } \nu_\alpha \nu_\beta = \nu_{\alpha+\beta}, \text{ falls } \alpha + \beta \text{ Wurzelform.}$$

[10] In **18** und **19** befinden wir uns auf den Fußspuren von F. Gantmakher, Canonical representations of automorphisms of a complex semi-simple group [Recueil Math. Moscou **5(47)** (1939), 101–146], insb. 128, 129. Da wir mehr mit dem Reellen rechnen müssen, verfahren wir in **18** ein bißchen anders als Gantmakher. Die Resultate von Gantmakher gehen (im Komplexen) weiter als unsere in **18** und **19**; doch spielt das für unsere Untersuchungen keine Rolle.

(3) $\nu_{\tilde{\alpha}}\overline{E_\alpha} = \nu_{\tilde{\alpha}}E_{\tilde{\alpha}} = AE_{\tilde{\alpha}} = A\overline{E_\alpha} = \overline{AE_\alpha} = \overline{\nu_\alpha E_\alpha} = \bar{\nu}_\alpha \bar{E}_\alpha$, also $\nu_{\tilde{\alpha}} = \overline{\nu_\alpha}$.

(1) und (2) kann man auch so interpretieren: ist $\alpha + \beta = 0$ bzw. $\alpha + \beta + \gamma = 0$ eine Relation zwischen Wurzelformen, so ist $\nu_\alpha \nu_\beta = 1$ bzw. $\nu_\alpha \nu_\beta \nu_\gamma = 1$. Nach 16 folgt hieraus:

$$\text{ist } \sum_\rho p_\rho \rho = 0, \qquad \text{so ist } \prod_\rho \nu_\rho^{p_\rho} = 1 \qquad (p_\rho \text{ ganz}).$$

Auf grund dessen können wir für eine beliebige lineare Kombination

$$\sigma = \sum_\rho q_\rho \rho \qquad (q_\rho \text{ ganz})$$

der Wurzelformen eindeutig definieren

$$\nu_\sigma = \prod_\rho \nu_\rho^{q_\rho}.$$

Diese σ besitzen eine Basis (für ganzzahlige lineare Kombination) $\sigma_1, \cdots, \sigma_l$. Definieren wir

$$\vartheta_{\sigma_\kappa} = \log \nu_{\sigma_\kappa} \text{ (irgendein Wert des log)}$$

und $$\vartheta_\sigma = \sum_\kappa r_\kappa \log \nu_{\sigma_\kappa} \quad \text{für} \quad \sigma = \sum_\kappa r_\kappa \sigma_\kappa !$$

Dann ist ϑ_σ linear in den σ und für Wurzelformen α

$$e^{\vartheta_\alpha} = \nu_\alpha .$$

Wir definieren weiter

$$\eta_\sigma = \tfrac{1}{2}(\vartheta_\sigma + \overline{\vartheta_{\tilde{\sigma}}}).$$

Dann ist auch η_σ linear in σ, reell für reelle σ und

$$e^{2\eta_\sigma} = \nu_\sigma \overline{\nu_{\tilde{\sigma}}} = \nu_\sigma^2, \qquad e^{\eta_\sigma} = \pm \nu_\sigma .$$

Bestimmen wir $H = H_\xi$ so, daß

$$-(\xi\sigma) = \eta_\sigma$$

wird, so ist H reell und erzeugt $h = \exp H$ einen reellen Automorphismus

$$B \ldots x \to h^{-1}xh$$

mit

$$h^{-1}E_\alpha h = \exp(\eta_\alpha) \cdot E_\alpha = \pm \nu_\alpha E_\alpha .$$

Der Automorphismus AB^{-1} von $\mathfrak{g}$ läßt $\mathfrak{h}$ elementweise fest und macht

$$AB^{-1}E_\alpha = \pm E_\alpha .$$

Es gibt nur endlich viel derartige Automorphismen (entsprechend der Zeichenwahl). Da $\mathfrak{g}$ zentrumfrei, a beliebig in $\mathfrak{h}^*$ und $h \in \mathfrak{h}$ ist, besitzt also $\mathfrak{h}$ in $\mathfrak{h}^*$ nur endlich viel Nebengruppen, w.z.b.w.

19. $\mathfrak{h}$ sei reguläre Untergruppe der halbeinfachen Gruppe $\mathfrak{g}$, $\mathfrak{h}'$ sei die größte Untergruppe von $\mathfrak{g}$, in der $\mathfrak{h}$ Normalteiler ist. Dann ist $\mathfrak{h}'/\mathfrak{h}$ endlich.

BEWEIS: Sei $a \in \mathfrak{h}'$. Der Automorphismus

$$A \cdots x \to a^{-1}xa$$

läßt $\mathfrak{h}$ als Ganzes fest. Wir setzen

$$AH_\lambda = H_{A\lambda}.$$

A wirkt dann linear auf die λ. Für eine Wurzelform ρ gilt

$$[H_\lambda E_\rho] = -(\rho, \lambda)E_\rho, \quad \text{also} \quad [H_{A\lambda}, AE_\rho] = -(\rho, \lambda)AE_\rho.$$

Also permutiert A die Wurzelformen. Ist A die identische Permutation der Wurzelformen, so ist

$$(\rho, A\lambda) = (\rho, \lambda) \text{ für alle } \rho,$$

also

$$A\lambda = \lambda;$$

dann läßt A also $\mathfrak{h}$ elementweise fest, und dann ist $a \in \mathfrak{h}^*$ (siehe 18). Da die Wurzelformen nur endlich viel Permutationen zulassen, besitzt $\mathfrak{h}^*$ in $\mathfrak{h}'$ nur endlich viel Nebengruppen, also ist $\mathfrak{h}'/\mathfrak{h}^*$ endlich, also nach 18 auch $\mathfrak{h}'/\mathfrak{h}$.

20. 1. Ist $\mathfrak{h}$ eine abelsche Gruppe, so sei $\mathfrak{h}^n$ die Gruppe aller h^n mit $h \in \mathfrak{h}$; der Durchschnitt aller $\mathfrak{h}^n$ ($n = 1, 2, \cdots$) heißt die *Verkürzung* von $\mathfrak{h}$.

2. Man zeigt leicht: Ist $\mathfrak{h}$ abelsch und $\mathfrak{h}_0$, die Komponente von 1 in $\mathfrak{h}$, Liesch, ist ferner $\mathfrak{h}/\mathfrak{h}_0$ eine Gruppe von endlich vielen Erzeugenden, so ist $\mathfrak{h}_0$ die Verkürzung von $\mathfrak{h}$.

3. Seien $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$ Liesche Untergruppen der Lieschen abelschen Gruppe $\mathfrak{h}$ und $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$ die zugehörigen Infinitesimalgruppen. Sei $\mathfrak{C} = \mathfrak{A} \cap \mathfrak{B}$. Dann ist $\mathfrak{c}$ die Verkürzung von $\mathfrak{a} \cap \mathfrak{b}$.

BEWEIS: Nach 20.2 brauchen wir nur zu zeigen: $(\mathfrak{a} \cap \mathfrak{b})/\mathfrak{c}$ besitzt endlich viel Erzeugende.—Sei

$$h \in \mathfrak{a} \cap \mathfrak{b}.$$

Dann

$$h = \exp H_\lambda, \qquad H_\lambda \in \mathfrak{A} \text{ (reell)},$$

(*)

$$= \exp H_\mu, \qquad H_\mu \in \mathfrak{B} \text{ (reell)};$$

also

(**)

$$(\rho, \lambda) = (\rho, \mu) + n_\rho \cdot 2\pi i \quad (n_\rho \text{ ganz})$$

für alle Wurzelformen ρ. Die Zahlen $n_\rho = n_\rho(h)$ fassen wir zu einem "Vektor" $n(h)$ zusammen. Der Vektor $n(h)$ ist durch h nicht eindeutig bestimmt, sondern nur mod $\mathbf{m}$, der additiven Gruppe aller Vektoren $n(1)$. Die $n(h)$ mod $\mathbf{m}$ bilden, wie man leicht sieht, eine additive Darstellung der Gruppe $\mathfrak{a} \cap \mathfrak{b}$. Ist $n(h) \equiv 0$ mod $\mathbf{m}$, so kann man λ und μ in (*) so bestimmen, daß $n(h) = 0$ wird; dann wird in (**) $(\rho, \lambda) = (\rho, \mu)$ für alle ρ, also $\lambda = \mu$, $H_\lambda = H_\mu \in \mathfrak{A} \cap \mathfrak{B} = \mathfrak{C}$ und $h \in \mathfrak{c}$.

Also ist $(\mathfrak{a} \cap \mathfrak{b})/\mathfrak{c}$ homomorphes Bild der (additiven) Gruppe der $n(h)$ mod $\mathbf{m}$, also sicher eine Gruppe von endlich vielen Erzeugenden, w.z.b.w.

21. Folgender Satz definiert rein gruppentheoretisch die regulären Untergruppen:

SATZ 2: *$\mathfrak{h}$ ist dann und nur dann reguläre Untergruppe der halbeinfachen Gruppe $\mathfrak{g}$, wenn gilt:*

1. *$\mathfrak{h}$ ist abelsch;*
2. *$\mathfrak{h}$ ist seine eigene Verkürzung;*
3. *$\mathfrak{h}$ ist maximal in bezug auf* 1 *und* 2 *zusammen;*
4. *ist $\mathfrak{h}_0$ irgendein Normalteiler von $\mathfrak{h}$ und $\mathfrak{h}/\mathfrak{h}_0$ abzählbar, ist weiter $\mathfrak{h}'(\subset\mathfrak{g})$ die größte Gruppe, in der $\mathfrak{h}_0$ Normalteiler ist, so ist auch $\mathfrak{h}'/\mathfrak{h}_0$ abzählbar.*

BEWEIS: Sei $\mathfrak{h}$ regulär. Dann ist 1 trivial, 2 ist in 20.2 und 3 ist in 18 und 20.2 enthalten. Nun zu 4: Sei $\mathfrak{h}/\mathfrak{h}_0$ abzählbar; dann[11] ist $\overline{\mathfrak{h}_0} = \mathfrak{h}$, also auch $\mathfrak{h}$ Normalteiler von $\mathfrak{h}'$ und nach 19 $\mathfrak{h}'/\mathfrak{h}$ endlich, also $\mathfrak{h}'/\mathfrak{h}_0$ abzählbar, w.z.b.w.

UMGEKEHRT: Mögen Voraussetzungen 1–4 gelten. Sei $a \in \overline{\mathfrak{h}}$. Nach 10 häufen sich die Komponenten von $\overline{\mathfrak{h}}$ nicht; da aber a Häufungspunkt von $\mathfrak{h}$ ist, gibt es ein $b \in \mathfrak{h}$, das in derselben Komponente von $\overline{\mathfrak{h}}$ wie a liegt. Nennen wir die Komponente von 1 in $\overline{\mathfrak{h}}$ nun $\mathfrak{h}_0$, so ist $b^{-1}a = c \in \mathfrak{h}_0$. Es ist $a = bc$, also $\overline{\mathfrak{h}} \subset \mathfrak{h}\cdot\mathfrak{h}_0$. Andererseits $\mathfrak{h} \subset \overline{\mathfrak{h}}$, $\mathfrak{h}_0 \subset \overline{\mathfrak{h}}$, also $\mathfrak{h}\cdot\mathfrak{h}_0 \subset \overline{\mathfrak{h}}$. Demnach $\overline{\mathfrak{h}} = \mathfrak{h}\cdot\mathfrak{h}_0$. Da $\mathfrak{h}$ nach Voraussetzung 1 abelsch ist, da ferner $\mathfrak{h}$ nach Voraussetzung 2 und $\mathfrak{h}_0$ als Liesche abelsche Gruppe (siehe 20.2) seine eigene Verkürzung ist, ist auch $\overline{\mathfrak{h}} = \mathfrak{h}\cdot\mathfrak{h}_0$ seine eigene Verkürzung, also nach Voraussetzung 3: $\mathfrak{h} = \overline{\mathfrak{h}}$. Also ist $\mathfrak{h}$ abgeschlossen.

Sei wieder $\mathfrak{h}_0$ die Komponente von 1 in $\mathfrak{h}$. Dann ist $\mathfrak{h}_0$ Liesch und $\mathfrak{h}/\mathfrak{h}_0$ abzählbar. Wenden wir auf dies $\mathfrak{h}_0$ nun Voraussetzung 4 an! Ist $\mathfrak{h}_0$ Normalteiler in $\mathfrak{h}'$, so muß $\mathfrak{h}'/\mathfrak{h}_0$ abzählbar sein. Insbesondere kann $\mathfrak{h}_0$ nun nicht Normalteiler einer Lieschen Gruppe höherer Dimension sein, ist also nach 14 regulär. Nach 18 (mit $\mathfrak{h}_0$ statt $\mathfrak{h}$ und $\mathfrak{h}$ statt $\mathfrak{h}^*$) und 20.2 ist $\mathfrak{h}_0$ die Verkürzung von $\mathfrak{h}$, also nach Voraussetzung 2: $\mathfrak{h}_0 = \mathfrak{h}$. Also ist $\mathfrak{h}$ regulär, w.z.b.w.

Definition weiterer Untergruppen

Eine feste reguläre Untergruppe $\mathfrak{h}$ werde zugrunde gelegt.

22. 1. $\mathfrak{e}_\alpha$ bzw. $\mathfrak{E}_\alpha$ heiße die von E_α und $E_{\bar{\alpha}}$ erzeugte (reelle) Untergruppe von $\mathfrak{g}$ bzw. $\mathfrak{G}$. Ist $\mathbf{A}$ eine Menge von Wurzelformen, so heiße $\mathfrak{e}_\mathbf{A}$ bzw. $\mathfrak{E}_\mathbf{A}$ die von der Gesamtheit der $\mathfrak{e}_\alpha$ bzw. $\mathfrak{E}_\alpha$ mit $\alpha \in \mathbf{A}$ erzeugte Untergruppe. Die Menge aller $\mathfrak{E}_\alpha$ heiße $\mathfrak{E}$, die Menge aller $\mathfrak{e}_\alpha$ heiße $\mathfrak{e}$ (α variabel).

2. Sei $\mathfrak{e} \in \mathfrak{e}$ und $\mathfrak{E}$ die zugehörige Infinitesimalgruppe, sei $\mathbf{A}$ die Menge aller α mit $E_\alpha \in \mathfrak{E}$ und $\mathbf{A}^*$ die Menge aller von $\mathbf{A}$ abhängigen Wurzelformen. Wir definieren $\mathfrak{e}^* = \mathfrak{e}_{\mathbf{A}^*}$ ($\mathfrak{e}^*$ hängt von $\mathfrak{e}$ ab); die Infinitesimalgruppe von $\mathfrak{e}^*$ heiße $\mathfrak{E}^*$. Nach 17 ist $\mathfrak{E}^*$ halbeinfach und $\mathfrak{E}^* \cap \mathfrak{H}$ reguläre Untergruppe von $\mathfrak{E}^*$.

---

[11] Die Überstreichung bedeutet im Augenblick den Übergang zur abgeschlossenen Hülle.

3. Sei **A** wie in 2 definiert. Die von den H_α mit $\alpha \in \mathbf{A}$ erzeugte (reelle) Gruppe heiße $\mathfrak{h}_\mathfrak{e}$, die zugehörige Infinitesimalgruppe $\mathfrak{H}_\mathfrak{e}$. Man bemerkt ohneweiteres: $\mathfrak{h}_\mathfrak{e} = \mathfrak{h}_{\mathfrak{e}^*}$.

4. Sei **A** wie in 2 definiert. Die von den auf **A** *senkrechten* H_λ erzeugte Gruppe heiße $\mathfrak{h}^\mathfrak{e}$, die zugehörige Infinitesimalgruppe $\mathfrak{H}^\mathfrak{e}$. Man sieht ohneweiteres: $\mathfrak{h}^\mathfrak{e} = \mathfrak{h}^{\mathfrak{e}^*}$. Ferner: $\mathfrak{H}^\mathfrak{e}$ und $\mathfrak{H}_\mathfrak{e}$ sind orthogonale Komplemente voneinander.

23. Folgender Satz definiert rein gruppentheoretisch die in 22 eingeführten Untergruppen.

SATZ 3: 1. *Die Untergruppe* $\mathfrak{e}$ *der halbeinfachen Gruppe* $\mathfrak{g}$ *gehört dann und nur dann zu* $\mathfrak{k}$, *wenn* $\mathfrak{e}$ *erzeugt wird von den Kommutatoren* $e^{-1}h^{-1}eh$ *mit* $e \in \mathfrak{e}$ *und* $h \in \mathfrak{h}$.

2. $\mathfrak{e}^*$ *ist die größte Gruppe aus* $\mathfrak{k}$ *mit* $\mathfrak{h}^\mathfrak{e} \perp \mathfrak{h}^{\mathfrak{e}^*}$.

3. $\mathfrak{h}_\mathfrak{e}$ *entsteht durch Verkürzung aus* $\mathfrak{e}^* \cap \mathfrak{h}$.

4. $\mathfrak{h}^\mathfrak{e}$ *entsteht durch Verkürzung aus der Gruppe* $(\mathfrak{h}^\mathfrak{e})'$ *aller* $h \in \mathfrak{h}$ *mit* $he = eh$ *für alle* $e \in \mathfrak{e}$.

Durch Satz 2 ist $\mathfrak{h}$ abstrakt definiert, Satz 3.1 definiert die Gesamtheit der $\mathfrak{e} \in \mathfrak{k}$ abstrakt, Satz 3.4 definiert dann abstrakt zu jedem $\mathfrak{e}$ das $\mathfrak{h}^\mathfrak{e}$, Satz 3.2 definiert abstrakt $\mathfrak{e}^*$ unter Verwendung von $\mathfrak{h}^\mathfrak{e}$, und schließlich wird durch Satz 3.3 $\mathfrak{h}_\mathfrak{e}$ aus $\mathfrak{e}^*$ und $\mathfrak{h}$ definiert.

24. Wir erbringen den Beweis von Satz 3 in der Reihenfolge 1, 4, 2, 3.

1. "Dann": Nach Anwendung 1 von Satz I (siehe 9) ist $\mathfrak{e}$ Liesch. Sei $\mathfrak{E}$ die zugehörige Infinitesimalgruppe; dann ist $[HX] \in \mathfrak{E}$ für alle $H \in \mathfrak{H}$ und $X \in \mathfrak{E}$. Schreiben wir nach 11.1

$$X = H + \sum \tau^\alpha E_\alpha ,$$

so erhalten wir

$$[H_\lambda X] = -\sum \tau^\alpha (\alpha\lambda) E_\alpha \in \mathfrak{E}.$$

Hieraus und aus der Realität von $\mathfrak{E}$ erschließt man leicht, daß $\mathfrak{E}$ von einer Menge von E_α, $E_{\bar{\alpha}}$ erzeugt wird.

"Nur dann": Wird $\mathfrak{E}$ von einer Menge von E_α, $E_{\bar{\alpha}}$ erzeugt, so ist es reell und so gilt $[HE] \in \mathfrak{E}$ für alle $E \in \mathfrak{E}$. Daraus folgt die Behauptung.

4. Da $\mathfrak{h}$ abelsch ist, existiert zu jedem $h \in \mathfrak{h}$ ein $H_\lambda \in \mathfrak{H}$ mit $h = \exp H_\lambda$. Es gilt

$$h^{-1}E_\alpha h = \exp(-(\lambda\alpha)) \cdot E_\alpha .$$

$h \in (\mathfrak{h}^\mathfrak{e})'$ ist also äquivalent mit

$$\exp(-(\lambda\alpha)) = 1$$

oder

$$(*) \qquad (\lambda\alpha) \equiv 0 \bmod 2\pi i$$

für alle $\alpha \in \mathbf{A}$ (Def. von **A**, siehe 22.2).

Sei $\mathfrak{K}$ die Gruppe aller (reellen) H_λ mit

$$(\lambda\alpha) = 0 \text{ für alle } \alpha \in \mathbf{A}.$$

Dann ist nach Definition (22.4)

$$(**) \qquad \mathfrak{K} = \mathfrak{H}^{\mathfrak{k}}.$$

Nach (*) erzeugt $\mathfrak{K}$ die Gruppe

$$\mathfrak{k} = \text{Komponente von } \iota \text{ in } (\mathfrak{h}^{\mathfrak{k}})'$$

und ist $(\mathfrak{h}^{\mathfrak{k}})'/\mathfrak{k}$ zyklisch. Nach 20.2 ist die Verkürzung von $(\mathfrak{h}^{\mathfrak{k}})'$ gleich $\mathfrak{k}$, ferner ist nach (**) $\mathfrak{k} = \mathfrak{h}^{\mathfrak{k}}$, also $\mathfrak{h}^{\mathfrak{k}}$ die Verkürzung von $(\mathfrak{h}^{\mathfrak{k}})'$, w.z.b.w.

2. Die Behauptung ergibt sich auf grund des Vorangehenden unmittelbar aus der Tatsache, daß

$$(\alpha\lambda) = 0 \text{ für alle } \alpha \in \mathbf{A}$$

und

$$(\alpha\lambda) = 0 \text{ für alle } \alpha \in \mathbf{A}^*$$

äquivalente Behauptungen sind.

3. Wir setzen $\mathfrak{k} = \mathfrak{e}^* \cap \mathfrak{h}$; $\mathfrak{k}_0$ sei die Komponente von ι in $\mathfrak{k}$, $\mathfrak{K}$ die zugehörige Infinitesimalgruppe. $\mathfrak{K}$ wird erzeugt von den $H_\alpha = [E_\alpha, E_{-\alpha}]$ mit $E_\alpha \in \mathfrak{E}^*$, also $\mathfrak{K} = \mathfrak{H}_{\mathfrak{e}^*} = \mathfrak{H}_{\mathfrak{e}}$, also

$$\mathfrak{k}_0 = \mathfrak{h}_{\mathfrak{e}}.$$

Nach 22.2 (Ende) ist $\mathfrak{k}_0$ reguläre Untergruppe der halbeinfachen Gruppe $\mathfrak{e}^*$, $\mathfrak{k} \supset \mathfrak{k}_0$, $\mathfrak{k}$ ist abelsch, also ist $\mathfrak{k}/\mathfrak{k}_0$ nach 18 endlich, also nach 20.2 $\mathfrak{k}_0$ $(= \mathfrak{h}_{\mathfrak{e}})$ die Verkürzung von $\mathfrak{k}$, w.z.b.w.

Die ausgezeichneten Untergruppen von $\mathfrak{h}$

25. $\mathfrak{M}$ sei die kleinste Menge von Untergruppen von $\mathfrak{H}$ mit der folgenden Eigenschaft:

1. $\mathfrak{M}$ enthält jedes $\mathfrak{H}_{\mathfrak{e}}$ und jedes $\mathfrak{H}^{\mathfrak{k}}$ mit $\mathfrak{e} \in \mathfrak{k}$.
2. $\mathfrak{M}$ enthält mit je zwei Gruppen ihren Durchschnitt und die von ihnen erzeugte Gruppe.

Von selber gilt dann nach 22.4 (und weil Durchschnitt und Erzeugung orthogonal-komplementäre Operationen sind):

3. $\mathfrak{M}$ enthält mit irgendeiner Gruppe auch ihr orthogonales Komplement.

$\mathfrak{e}_\alpha$ war die von E_α und $E_{\bar\alpha}$ erzeugte Gruppe. Statt $\mathfrak{H}^{\mathfrak{k}_\alpha}$ bzw. $\mathfrak{H}_{\mathfrak{e}_\alpha}$ wollen wir kurz $\mathfrak{H}^\alpha$ bzw. $\mathfrak{H}_\alpha$ schreiben. Analog $\mathfrak{h}_\alpha$, $\mathfrak{h}^\alpha$. $\mathfrak{H}_\alpha$ wird erzeugt von H_α und $H_{\bar\alpha}$. Allgemeiner sei (bei beliebigem λ) $\mathfrak{H}_\lambda$ die von H_λ und $H_{\bar\lambda}$ erzeugte Gruppe; ferner $\mathfrak{H}^\lambda$ das orthogonale Komplement von $\mathfrak{H}_\lambda$. Analog $\mathfrak{h}_\lambda$, $\mathfrak{h}^\lambda$.

Wir werden in Zukunft häufig die Fälle unterscheiden:

λ reell (d.h. *rein* reell),

λ imaginär (d.h. *rein* imaginär),

λ komplex (d.h. weder rein reell, noch rein imaginär).
In den beiden ersten Fällen ist $\mathfrak{H}_\lambda$ eindimensional, im dritten Fall zweidimensional (entsprechend $\mathfrak{H}^\lambda$ $(n-1)$- oder $(n-2)$-dimensional).

Wir bemerken schließlich:

4. Enthält $\mathfrak{M}$ eine gewisse nichtisotrope (11.2) eindimensionale Untergruppe $\mathfrak{H}_\lambda$ von $\mathfrak{H}_\alpha$, so enthält es auch noch eine zweite eindimensionale Untergruppe $\mathfrak{H}_\mu$ von $\mathfrak{H}_\alpha$ (nämlich das orthogonale Komplement von $\mathfrak{H}_\lambda$ schneidet $\mathfrak{H}_\alpha$ in einer solchen: $\mu = p\alpha + \bar{p}\bar{\alpha}$, $(\lambda, \mu) = 0$, also wegen der Nichtisotropie $\lambda \neq \mu$).

26. Die Menge $\mathfrak{m}$ (von Untergruppen von $\mathfrak{h}$) sei folgendermaßen definiert: $\mathfrak{k} \in \mathfrak{m}$ dann und nur dann, wenn $\mathfrak{k}$ die von einer Gruppe $\mathfrak{K} \in \mathfrak{M}$ erzeugte Liesche Gruppe ist.

$\mathfrak{m}$ ist dann auch die kleinste Menge von Untergruppen von $\mathfrak{h}$ mit der folgenden Eingenschaft:

1. $\mathfrak{m}$ enthält jedes $\mathfrak{h}_\mathfrak{e}$ und jedes $\mathfrak{h}^\mathfrak{e}$ mit $\mathfrak{e} \in \mathfrak{k}$.

2. $\mathfrak{m}$ enthält mit je zwei Gruppen die Verkürzung ihres Durchschnitts und die von ihnen erzeugte Gruppe. Um das einzusehen, beachte man 20.3.

Nach Satz 3 konnten wir die $\mathfrak{h}_\mathfrak{e}$ und $\mathfrak{h}^\mathfrak{e}$ rein gruppentheoretisch definieren. Somit ist auch die Menge $\mathfrak{m}$ (als Funktion von $\mathfrak{h}$ natürlich) rein gruppentheoretisch definiert.

27. SATZ 4: *Sei $\mathfrak{g}$ halbeinfach und von zweiter Art. Dann gilt*: 1. *Ist die Gruppe $\mathfrak{h}_\mathfrak{e}$ eine minimale*[12] *Gruppe aus der Menge $\mathfrak{m}$, so ist $\mathfrak{h}_\mathfrak{e}$ gleich einem $\mathfrak{h}_\alpha$ und α reelle oder imaginäre Wurzelform.* 2. *Sind alle Wurzelformen komplex, so gibt es Wurzelformen α und β, die auch gleich sein können, derart daß $\mathfrak{h}_{\alpha+\bar{\alpha}}$ und $\mathfrak{h}_{\beta-\bar{\beta}}$ Elemente von $\mathfrak{m}$ sind.*

BEWEIS: Daß ein minimales $\mathfrak{h}_\mathfrak{e}$ ein $\mathfrak{h}_\alpha$ ist, ist klar. Die Annahme, α sei komplex, führen wir zum Widerspruch.

Klar ist, daß

$$(\alpha, \bar{\alpha}) = 0$$

sein müßte, da sonst nach 11.3 entweder $\alpha + \bar{\alpha}$ oder $\alpha - \bar{\alpha}$ Wurzelform, also $\mathfrak{h}_{\alpha+\bar{\alpha}}$ oder $\mathfrak{h}_{\alpha-\bar{\alpha}}$ in $\mathfrak{m}$, also $\mathfrak{h}_\alpha$ nicht minimal wäre.

Voraussetzungsgemäß ist $\mathfrak{g}$ komplex-einfach, also gibt es nach 12 eine Wurzelform β mit

$$(\alpha, \beta) \neq 0, \qquad (\bar{\alpha}, \beta) \neq 0.$$

Dann ist auch

$$(\bar{\alpha}, \bar{\beta}) \neq 0, \qquad (\alpha, \bar{\beta}) \neq 0.$$

Wir unterscheiden drei Fälle:

A: β ist reell oder imaginär.—Wegen $(\alpha, \beta) \neq 0$ ist $\mathfrak{H}_\alpha$ nicht senkrecht auf $\mathfrak{H}_\beta$, also (siehe 22.4, Ende) $\mathfrak{H}_\alpha \not\subset \mathfrak{H}^\beta$. Da ferner $\mathfrak{H}^\beta$ $(l-1)$-dim. ist, ist $\mathfrak{H}_\alpha \cap \mathfrak{H}^\beta$

[12] Natürlich minimal bei Vernachlässigung der Gruppe (I).

eindimensional, also eine zu $\mathfrak{M}$ gehörige echte Untergruppe von $\mathfrak{H}_\alpha$. Widerspruch zur Minimalität von $\mathfrak{h}_\alpha$!

B: β ist komplex und

$$(\beta, \bar{\beta}) \neq 0.$$

Nach 11.3 ist etwa

$$\gamma = \beta + \bar{\beta}$$

Wurzelform (denn Fall, daß $\beta - \bar{\beta}$ Wurzelform ist, behandelt man ganz analog), also $\mathfrak{H}_\gamma \in \mathfrak{M}$. Als Wurzelvektor ist H_γ nichtisotrop, also existiert nach 24.4 noch ein $\mathfrak{H}_\delta \in \mathfrak{M}$, $\mathfrak{H}_\delta \subset \mathfrak{H}_\beta$, $\mathfrak{H}_\delta \neq \mathfrak{H}_\gamma$ (δ braucht nicht Wurzelform zu sein). $\mathfrak{H}_\gamma$ und $\mathfrak{H}_\delta$ erzeugen zusammen $\mathfrak{H}_\beta$, und da $(\alpha, \beta) \neq 0$ war, können nicht beide auf $\mathfrak{H}_\alpha$ senkrecht sein. Es gibt demnach ein eindimensionales $\mathfrak{H}_\epsilon$ ($\epsilon = \gamma$ oder δ) in $\mathfrak{M}$, das nicht auf $\mathfrak{H}_\alpha$ senkrecht ist. Das orthogonale Komplement $\mathfrak{H}'$ von $\mathfrak{H}_\epsilon$ gehört nach 25.3 zu $\mathfrak{M}$ und schneidet $\mathfrak{H}_\alpha$ eindimensional. $\mathfrak{H}' \cap \mathfrak{H}_\alpha \in \mathfrak{M}$ im Widerspruch zur Minimalität von $\mathfrak{h}_\alpha$!

C: β ist komplex und

$$(\beta, \bar{\beta}) = 0.$$

Wegen $(\alpha, \beta) \neq 0$ ist $\alpha + \beta$ oder $\alpha - \beta$ Wurzelform. Wir dürfen annehmen, $\alpha + \beta$ (sonst ersetzen wir β durch $-\beta$).

Wegen $(\alpha, \bar{\alpha}) = 0$ haben wir

$$(\alpha + \beta, \bar{\alpha}) = (\beta, \bar{\alpha}) \neq 0.$$

Also ist $\alpha + \beta \pm \bar{\alpha}$ Wurzelform ($\pm$ ist im Folgenden als "+ *oder* −" zu lesen). Wir unterscheiden zwei Unterfälle:

I. Sei

$$(\alpha + \beta \pm \bar{\alpha}, \bar{\beta}) = 0$$

Dann (wegen $(\beta, \bar{\beta}) = 0$):

$$(*) \qquad \begin{aligned} (\alpha \pm \bar{\alpha}, \bar{\beta}) &= 0, \\ (\alpha \pm \bar{\alpha}, \beta) &= 0. \end{aligned}$$

Also (22.4, Ende)

$$\mathfrak{H}^\beta \supset \mathfrak{H}_{\alpha \pm \bar{\alpha}}.$$

Andererseits wegen $(\alpha, \beta) \neq 0$

$$\mathfrak{H}^\beta \not\supset \mathfrak{H}_\alpha.$$

Zusammen:

$$(\mathrm{CI}, 1) \qquad \mathfrak{H}^\beta \cap \mathfrak{H}_\alpha = \mathfrak{H}_{\alpha \pm \bar{\alpha}} \in \mathfrak{M}.$$

Widerspruch zur Minimalität von $\mathfrak{h}_\alpha$.

Wir schließen aber noch weiter, falls alle Wurzelformen komplex sind: $(\alpha + \beta, \overline{\alpha + \beta}) = 0$, also $(\alpha, \bar{\beta}) = -(\bar{\alpha}, \beta)$. Einsetzen in (*) liefert:

$$(\alpha, \beta \mp \bar{\beta}) = 0,$$

$$(\bar{\alpha}, \beta \mp \bar{\beta}) = 0.$$

Hieraus schließt man wie vorhin:

(CI, 2) $$\mathfrak{H}^{\alpha} \cap \mathfrak{H}_{\beta} = \mathfrak{H}_{\beta \mp \bar{\beta}} \in \mathfrak{M}.$$

II. Sei nun

$$(\alpha + \beta \pm \bar{\alpha}, \bar{\beta}) \neq 0.$$

Dann hat man vier Möglichkeiten:

a) $\alpha + \beta + \bar{\alpha} + \bar{\beta}$ Wurzelform

b) $\alpha + \beta + \bar{\alpha} - \bar{\beta}$ " "

a') $\alpha + \beta - \bar{\alpha} - \bar{\beta}$ " "

b') $\alpha + \beta - \bar{\alpha} + \bar{\beta}$ " " .

Wir behandeln nur a und b, da a' und b' genauso verlaufen.

a) $\gamma = \alpha + \bar{\alpha} + \beta + \bar{\beta}$ ist (reelle) Wurzelform. $\mathfrak{H}_{\gamma} \subset \mathfrak{H}_{\alpha+\beta}$, also existiert nach 25.4 noch ein $\mathfrak{H}_{\delta} \in \mathfrak{M}$, $\mathfrak{H}_{\delta} \subset \mathfrak{H}_{\alpha+\beta}$, $\mathfrak{H}_{\delta} \neq \mathfrak{H}_{\gamma}$. $\mathfrak{H}_{\gamma}$ und $\mathfrak{H}_{\delta}$ erzeugen zusammen $\mathfrak{H}_{\alpha+\beta}$, und da $(\alpha + \beta, \bar{\alpha}) = (\beta, \bar{\alpha}) \neq 0$ war, können nicht beide auf $\mathfrak{H}_{\alpha}$ senkrecht sein. Sind $\mathfrak{H}^{\gamma}$ und $\mathfrak{H}^{\delta}$ die orthogonalen Komplemente, so gilt demnach

$$\mathfrak{H}^{\gamma} \supset \mathfrak{H}_{\alpha} \quad \text{oder} \quad \mathfrak{H}^{\delta} \supset \mathfrak{H}_{\alpha},$$

also ist $\mathfrak{H}^{\gamma} \cap \mathfrak{H}_{\alpha}$ oder $\mathfrak{H}^{\delta} \cap \mathfrak{H}_{\alpha}$ eindimensional und in $\mathfrak{M}$ (siehe 25.4) im Widerspruch zur Minimalität von $\mathfrak{h}_{\alpha}$.

b) $\gamma = \alpha + \bar{\alpha} + \beta - \bar{\beta}$ ist Wurzelform, $2\Re\alpha = \Re\gamma$, $2\Im\beta = \Im\gamma$. $\mathfrak{H}_{\gamma} \cap \mathfrak{H}_{\alpha} = \mathfrak{H}_{\alpha+\bar{\alpha}} \in \mathfrak{M}$, $\mathfrak{H}_{\gamma} \cap \mathfrak{H}_{\beta} = \mathfrak{H}_{\beta-\bar{\beta}} \in \mathfrak{M}$.—Widerspruch zur Minimalität von $\mathfrak{h}_{\alpha}$.

Damit ist die erste Hälfte des Satzes bewiesen. Die Voraussetzung der zweiten Hälfte (alle Wurzelformen komplex) hat auch wieder $(\alpha, \bar{\alpha}) = 0$ zur Folge; sie hat weiter zur Folge, daß wir uns in den Fällen CI oder CIIb befinden —in beiden Fällen haben wir aber die zweite Behauptung bereits bewiesen (siehe die Formeln CI, 1 und CI, 2 und die Formeln in CIIb).

28. Ist

$$(\alpha, \bar{\alpha}) = 0,$$

so ist

$$[H'_{\alpha}, E'_{\alpha}] = \vartheta E'_{\alpha} \qquad (\vartheta \text{ reell}),$$

falls

$$E'_{\alpha} = E_{\alpha} + E_{\bar{\alpha}}, \qquad H'_{\alpha} = H_{\alpha} + H_{\bar{\alpha}}, \qquad \vartheta = (\alpha, \alpha).$$

ist.

Rein gruppentheoretische Definition kompakter Untergruppenkeime

29. SATZ 5: *$\mathfrak{g}$ sei halbeinfach von zweiter Art. Man wähle, wenn möglich, ein $\mathfrak{h}' = \mathfrak{h}_t$, so daß es minimal in $\mathfrak{m}$ ist, und dabei $\mathfrak{e} \in \mathfrak{k}$ (auch $\mathfrak{e}'$ genannt) so, daß es auch minimal ist. Ist das nicht möglich, so wähle man $\mathfrak{e} \in \mathfrak{k}$ und eine in $\mathfrak{m}$ minimale Untergruppe $\mathfrak{h}'$ von $\mathfrak{h}_t$, so daß gilt:*

1. *$\mathfrak{e}$ ist minimal;*

2. *es gibt eine echte Untergruppe $\mathfrak{e}' \neq (1)$ von $\mathfrak{e}$, derart daß die Kommutatoren $h^{-1}e^{-1}he$ $(h \in \mathfrak{h}', e \in \mathfrak{e}')$ die Gruppe $\mathfrak{e}'$ erzeugen.*

Diese Wahl ist möglich, und es gilt weiter:

A. *Entweder ist $\mathfrak{h}'$ eine kompakte Gruppe.*

B. *Oder $\mathfrak{h}'$ ist nicht kompakt,*

$$\mathfrak{h}' = (\exp tH'), \qquad \mathfrak{e}' = (\exp tE'),$$

$$[H'E'] = \vartheta E' \qquad (\vartheta \text{ reell}).$$

Die beiden Fälle sind noch abstrakt dadurch unterschieden, daß in A *$\mathfrak{h}'$ Elemente endlicher Ordnung besitzt, in* B *nicht.*

30. BEWEIS VON SATZ 5: Sind nicht alle Wurzelformen komplex, so ist $\mathfrak{h}'$ wegen der Minimalität und nach Satz 4 ein $\mathfrak{h}_\alpha$ mit reellem oder imaginärem α; $\mathfrak{e}' = \mathfrak{e}_\alpha$ oder $\mathfrak{e}_{-\alpha}$. Ist α reell, so befinden wir uns im Fall B, $\vartheta = (\alpha, \alpha)$. Sei α imaginär. Dann wird $\mathfrak{e}_\alpha$ erzeugt von E_α, $E_{-\alpha}$. Eine Basis von $\mathfrak{E}_\alpha$ ist E_α, $E_{-\alpha}$, H_α, und es gilt

$$[E_\alpha, E_{-\alpha}] = H_\alpha,$$

also nach Übergang zum Konjugierten

$$[E_{-\alpha}, E_\alpha] = \bar{H}_\alpha,$$

also $\bar{H}_\alpha = -H_\alpha$, also H_α imaginär. Wegen $(\alpha, \alpha) < 0$ und 13 ist auch (α, β) reell für alle Wurzelformen α, β. Ferner ist

$$H' = iH_\alpha$$

reell und

$$\mathfrak{h}' = (\exp tH')_{t \text{ reell}}.$$

In

$$[H'E_\beta] = -i(\alpha, \beta)E_\beta$$

sind wegen der Realität der (α, β) die Koeffizienten $-i(\alpha, \beta)$ imaginär. Also sind alle Wurzeln von H' imaginär und rationale Vielfache voneinander, und demnach ist $\mathfrak{h}'$ kompakt (Fall A).

Sind hingegen alle Wurzelformen komplex, so versagt nach Satz 4, 1 die erste Bestimmungsmethode für $\mathfrak{h}'$. Nach Satz 4, 2 und nach 28 ist die vorgeschriebene Wahl von $\mathfrak{h}'$ und $\mathfrak{e}'$ möglich. Sei diese Wahl irgendwie geschehen. Dann ist $\mathfrak{h}'$ eindimensional nach Satz 4, 1 und $\mathfrak{e}'$ Liesch nach Satz 1 und eindimensional als echte Untergruppe von $\mathfrak{e}$. Wir befinden uns somit wieder im Fall B.

31. SATZ 6: *Ein eindimensionaler kompakter Untergruppenkeim der halbeinfachen Gruppe zweiter Art* $\mathfrak{g}$ *läßt sich rein gruppentheoretisch beschreiben.*

BESCHREIBUNG: Im Falle A sind wir nach Satz 5 bereits fertig.—Fall B:

$$[H'E'] = \vartheta E',$$

$$h_t = \exp\left(\frac{t}{\vartheta} H'\right), \qquad e_s = \exp(sE'),$$

$$h_t E' h_t^{-1} = (\exp t) \cdot E'$$

$$(*) \qquad h_t e_s h_t^{-1} = e_{s \cdot \exp t}.$$

Sei $e_{s_0} (\neq 1)$ beliebig in $\mathfrak{e}'$; die Menge aller $h_t e_{s_0} h_t^{-1}$ nennen wir einen Strahl. Nach (*) gibt es in $\mathfrak{e}'$ nur die zwei Strahlen e_τ, $\tau > 0$, und e_τ, $\tau < 0$. Sei $\mathfrak{s}$ ein Strahl von $\mathfrak{e}'$, $e_{s_0} \in \mathfrak{s}$ und $\mathfrak{t}$ der andere Strahl.

$$\mathfrak{u} = \mathfrak{s} \cap e_{s_0} \cdot \mathfrak{t}$$

ist die Menge aller e_s mit s zwischen 0 und s_0.

$$\mathfrak{u} \cup \mathfrak{u}^{-1} \cup (e_{s_0}) \cup (e_{s_0}^{-1}) \cup (1)$$

ist also ein kompakter Keim von $\mathfrak{e}'$ und rein gruppentheoretisch beschrieben.

Rein gruppentheoretische Beschreibung einer kompakten Umgebung der Eins

32. Die Elemente

$$X_1, \cdots, X_{p_1} \qquad (p_1 > 1)$$

mögen die r-parametrige Gruppe $\mathfrak{G}$ erzeugen (Halbeinfachheit brauchen wir hier nicht vorauszusetzen). Man erweitere dies System zu einem System

$$X_1, \cdots, X_{p_1}, \cdots, X_{p_2}$$

durch Hinzufügung der Kommutatoren

$$[X_\rho X_\sigma] \qquad (\rho, \sigma \leqq p_1),$$

dies analog zu einem System

$$X_1, \cdots, X_{p_1}, \cdots, X_{p_2}, \cdots, X_{p_3}$$

durch Hinzufügung der Kommutatoren

$$[X_\rho X_\sigma] \qquad (\rho, \sigma \leqq p_2),$$

und fahre sofort. Dann gibt es ein u, das nur von r abhängt, derart daß

$$X_1, \cdots, X_u$$

eine Basis von $\mathfrak{G}$ bilden.

Man nehme nämlich $u = p_{r-1}$. Ist v so bestimmt, daß die

$$X_1, \cdots, X_{p_v}, \cdots, X_{p_{v+1}} \qquad (X_{v+1})$$

von den

$$X_1, \cdots, X_{p_\nu} \qquad (\mathfrak{X}_\nu)$$

linear abhängen, so enthält $\mathfrak{X}_\nu$ eine Basis von $\mathfrak{G}$. Sei ν minimal. Dann liegt für $\varrho < \nu$ in

$$X_1, \cdots, X_{p_\varrho}, \cdots, X_{p_\varrho+1}$$

mindestens ein linear unabhängiges Element mehr als in

$$X_1, \cdots, X_{p_\varrho}.$$

Da es in $\mathfrak{G}$ nur r unabhängige Elemente gibt, muß also $\nu < r$ sein. Somit reicht $u = p_{r-1}$ in der Tat aus.

33. SATZ 7: *In der einfachen Gruppe zweiter Art $\mathfrak{g}$ läßt sich eine kompakte Teilmenge U, die ι als inneren Punkt enthält, rein gruppentheoretisch beschreiben.*

BESCHREIBUNG: $\mathfrak{k}$ sei der kompakte eindimensionale Untergruppenkeim aus Satz 6. Wegen der Einfachheit von $\mathfrak{g}$ kann man $a_1, \cdots, a_p$ so wählen, daß die ebenfalls kompakten Keime

$$\mathfrak{k}_\nu = a_\nu^{-1}\mathfrak{k}a_\nu \qquad (\nu = 1, \cdots, p)$$

$\mathfrak{g}$ erzeugen;

$$\mathfrak{k}_\nu : k_\nu(s) \qquad (0 \leqq s \leqq 1).$$

Zu den $k_\nu(s)$ füge man die Kommutatorkurven

$$k_\varrho(s)^{-1}k_\sigma(s)^{-1}k_\varrho(s)k_\sigma(s) \qquad (\varrho < \sigma),$$

so daß das System

$$k_1(s), \cdots, k_{p_1}(s), \cdots, k_{p_2}(s) \qquad (p_1 = p)$$

entsteht. So fahre man fort bis zum Index u aus 32. Nach 32 kommt unter den Hauptteilen von

$$k_1(s), \cdots, k_u(s)$$

eine Basis von $\mathfrak{G}$ vor. Nach 4 enthält die offenbar kompakte Menge

$$U : k_1(s_1) \cdots k_u(s_u) \qquad (0 \leqq s_\nu \leqq 1)$$

eine Umgebung von ι, wie wir es wünschten.

Beweis des Hauptsatzes I

34. SATZ 8: *In der einfachen Gruppe zweiter Art $\mathfrak{g}$ lassen sich beliebig kleine kompakte Teilmengen V, die ι als inneren Punkt enthalten, rein gruppentheoretisch beschreiben.*

BESCHREIBUNG: a sei beliebig $\neq \iota$. U sei gemäß Satz 7 gewählt.

$$V = \bigcup_{\substack{b_\nu \in U \\ c_\nu \in U}} \prod_{\nu=1}^{r} c_\nu^{-1}(b_\nu^{-1}a^{-1}b_\nu a)c_\nu.$$

Wörtlich wie bei v. d. Waerden, a.a.O.,[2] 783–784, zeigt man, daß die (von a abhängenden) V die gewünschten Eigenschaften besitzen. Nur verwendet man hier statt der Kompaktheit von $\mathfrak{g}$ die von U (Satz 7).

35. Der Satz 8 ist mit unserm Hauptsatz I äquivalent (siehe dazu auch v. d. Waerden, a.a.O.,[2] 784–785), wenn wir *beide* miteinander isomorphe Gruppen als von zweiter Art annehmen. Da die Einfachheit von selber ein rein gruppentheoretischer Begriff ist, haben wir nur noch zu zeigen, daß auch die "zweite Art" oder—was auf dasselbe hinauskommt—die "erste Art" rein gruppentheoretisch zu beschreiben ist, um den Hauptsatz I ganz bewiesen zu haben. Das geschieht nun:

Rein gruppentheoretische Beschreibung der "ersten" Art

36. Satz 4′: *Eine halbeinfache Gruppe ist dann und nur dann von erster Art, wenn für je zwei minimale Gruppen* $\mathfrak{e}$ *und* $\mathfrak{e}'$ *aus* $\mathfrak{k}$ *gilt*:

1) $\mathfrak{e}$ *ist abelsch*;
2) $\mathfrak{e}$ *enthält Elemente endlicher Ordnung*;
3) *Verkürzung von* $\mathfrak{h}_\mathfrak{e} \cap \mathfrak{h}^{\mathfrak{e}'} = (1)$ *oder* $\mathfrak{h}_\mathfrak{e}$;
4) *Verkürzung von* $\mathfrak{h}_\mathfrak{e} \cap \mathfrak{h}_{\mathfrak{e}'} = (1)$ *oder* $\mathfrak{h}_\mathfrak{e}$.

Beweis: "Nur dann": Die Wurzelformen einer Gruppe erster Art zerfallen in zwei Systeme $\mathbf{P}$ und $\bar{\mathbf{P}}$. Ist $\alpha \in \mathbf{P}$, so ist $\bar{\alpha} \in \bar{\mathbf{P}}$ und umgekehrt; ist $\alpha, \beta \in \mathbf{P}$, so ist $(\alpha, \bar{\beta}) = 0$.

Sei $\mathfrak{g}$ von erster Art. Jedes minimale $\mathfrak{e}$ ist von der Gestalt $\mathfrak{e}_\sigma$. Das zugehörige $\mathfrak{E}_\alpha$ wird von E_α und $E_{\bar{\alpha}}$ erzeugt. Wegen $(\alpha, \bar{\alpha}) = 0$ ist $[E_\alpha E_{\bar{\alpha}}] = 0$, also 1 erfüllt.

Seien $\alpha, \rho \in \mathbf{P}$. Man bilde

$$K = i(\bar{\alpha}, \bar{\alpha})H_\alpha - i(\alpha, \alpha)H_{\bar{\alpha}};$$

aus $(\alpha, \bar{\rho}) = (\bar{\alpha}, \rho) = 0$ folgt

$$[KE_\rho] = -i(\bar{\alpha}, \bar{\alpha})(\alpha, \rho)E_\rho,$$

$$[KE_{\bar{\rho}}] = \quad i(\alpha, \alpha)(\bar{\alpha}, \bar{\rho})E_{\bar{\rho}}.$$

Aus 11.3 und $(\alpha, \alpha) = (\bar{\alpha}, \bar{\alpha}) < 0$ schließt man, daß die Koeffizienten rechts (für alle $\rho \in \mathbf{P}$), d.h. die Wurzeln von K, rein imaginär und rationale Vielfache voneinander sind. Die von K erzeugte Untergruppe $\mathfrak{k}$ von $\mathfrak{h}_\mathfrak{e}$ ist also kompakt, und daraus folgt 2.

$\mathfrak{e} = \mathfrak{e}_\alpha$, $\mathfrak{e}' = \mathfrak{e}_\beta$. Sei $H_\lambda \in \mathfrak{H}_\alpha \cap \mathfrak{H}^\beta$. Dann

$$\lambda = p\alpha + \bar{p}\bar{\alpha}, \qquad (\lambda, \beta) = 0,$$

also wegen $(\bar{\alpha}, \beta) = 0$

$$p(\alpha, \beta) = 0.$$

Also entweder $(\alpha, \beta) = 0$ und dann $\mathfrak{H}_\alpha$ senkrecht auf $\mathfrak{H}_\beta$ und $\mathfrak{H}_\alpha \cap \mathfrak{H}^\beta = \mathfrak{H}_\alpha$. Oder $p = 0$, $\bar{p} = 0$, $\lambda = 0$ und dann $\mathfrak{H}_\alpha \cap \mathfrak{H}^\beta = (0)$. Also gilt 3 (siehe 20.3).

Wir betrachten nun $H_\lambda \in \mathfrak{H}_\alpha \cap \mathfrak{H}_\beta$.

$$\lambda = p\alpha + \bar{p}\bar{\alpha} = q\beta + \bar{q}\bar{\beta}$$

"Multiplikation" mit ρ liefert:

$$p(\alpha\rho) = q(\beta\rho) \qquad \text{für alle } \rho,$$

also

$$p\alpha = q\beta.$$

Also entweder $p = q = 0$, $\lambda = 0$ und dann $\mathfrak{H}_\alpha \cap \mathfrak{H}_\beta = (0)$. Oder $p \neq 0$ und dann $\mathfrak{H}_\alpha = \mathfrak{H}_\beta$. Also gilt 4 (siehe 20.3).

"Dann": Wegen Voraussetzung 1 kann das α eines minimalen $\mathfrak{e}_\alpha$ nicht imaginär sein ($[E_\alpha E_{-\alpha}]$ ist doch $= H_\alpha$), wegen Voraussetzung 2 kann es nicht reell sein ($\mathfrak{e}_\alpha$ ist dann offen); also ist für ein minimales $\mathfrak{e}_\alpha$ das α komplex. Da andererseits $\mathfrak{e}_\alpha$ für reelles oder imaginäres α sicher minimal ist, müssen alle Wurzelformen komplex sein. Wäre $\mathfrak{g}$ von zweiter Art, so kämen nur die Fälle CI und CIIb aus 27 in Frage. CI ist ausgeschlossen wegen $\mathfrak{H}^\alpha \cap \mathfrak{H}_\beta \neq \mathfrak{H}_\beta$, $\neq(0)$ (Formel CI, 1) und CIIb wegen $\mathfrak{H}_\gamma \cap \mathfrak{H}_\alpha \neq \mathfrak{H}_\alpha$, $\neq(0)$. Also ist $\mathfrak{g}$ von erster Art.

Einfache Gruppen mit diskretem Zentrum

37. Wir beweisen die Bemerkung 1 zu Hauptsatz I.

Die einfachen Gruppen zweiter Art $\mathfrak{f}$ und $\mathfrak{g}$ dürfen ein diskretes Zentrum besitzen. A sei ein Isomorphimus, $A\mathfrak{f} = \mathfrak{g}$. Die resp. Faktorgruppen nach dem Zentrum seien $\mathfrak{f}'$ und $\mathfrak{g}'$; φ und ψ seien die zugehörigen Homomorphismen, $\varphi\mathfrak{f} = \mathfrak{f}'$, $\psi\mathfrak{g} = \mathfrak{g}'$. A induziert einen Isomorphismus A', $A'\mathfrak{f}' = \mathfrak{g}'$,

$$A'\varphi = \psi A. \tag{1}$$

Die universellen Überlagerungsgruppen von $\mathfrak{f}$ und $\mathfrak{g}$ seien $\tilde{\mathfrak{f}}$ und $\tilde{\mathfrak{g}}$, die zugehörigen Projektionen Φ und Ψ, $\Phi\tilde{\mathfrak{f}} = \mathfrak{f}$, $\Psi\tilde{\mathfrak{g}} = \mathfrak{g}$. A' ist nach Hauptsatz I stetig und induziert, wie man weiß, einen stetigen Homomorphismus $\tilde{A}$, $\tilde{A}\tilde{\mathfrak{f}} = \tilde{\mathfrak{g}}$,

$$A'\varphi\Phi = \psi\Psi\tilde{A}. \tag{2}$$

Wir definieren

$$\begin{aligned} A\Phi &= B, \\ \Psi\tilde{A} &= C. \end{aligned} \tag{3}$$

Aus (1), (2), (3) folgt

$$\psi B = \psi A\Phi = A'\varphi\Phi = \psi\Psi\tilde{A} = \psi C. \tag{4}$$

Das ψ-Urbild von $\mathfrak{1}$ ist das Zentrum von $\mathfrak{g}$, also folgt aus (4):

Für jedes $x \in \tilde{\mathfrak{f}}$ ist $(Bx)\cdot(Cx)^{-1}$ im Zentrum von $\mathfrak{g}$. Da das Zentrum abzählbar ist, kann der Normalteiler $\mathfrak{d}$ aller x mit

$$Bx = Cx$$

nicht abzählbar sein, sicher also nicht im Zentrum von $\mathfrak{f}$ liegen. Nach Anwendung 4 von Satz 1 ist dann $\mathfrak{d} = \mathfrak{f}$, also

$$Bx = Cx$$

für alle x. Da $\tilde{A}$, also nach (3) auch C stetig war, ist B stetig. Da Φ als im Kleinen topologische Abbildung beiderseits stetig ist (Bilder offener Mengen sind offen), ist nach (3) auch A stetig, w.z.b.w.

AMSTERDAM, HOLLAND

ANNALS OF MATHEMATICS
Vol. 42, No. 5, December, 1941

A TAUBERIAN THEOREM FOR PARTITIONS

BY A. E. INGHAM

(Received August 11, 1940)

1. Introduction

The formulae

$$(1) \qquad p(n) \sim \frac{e^{\pi\sqrt{(2n/3)}}}{(4\sqrt{3})n}, \qquad q(n) \sim \frac{e^{\pi\sqrt{(n/3)}}}{4\cdot 3^{\frac{1}{4}}\cdot n^{\frac{3}{4}}}, \qquad (n \to \infty),$$

where $p(n)$ is the number of unrestricted partitions of n and $q(n)$ the number of partitions of n into unequal parts (or into odd parts), were discovered by Hardy and Ramanujan [4], and independently by Uspensky [7]. Their proofs give much more than simple asymptotic equality, but demand a rather special study of the generating functions

$$\sum_{n=0}^{\infty} p(n)z^n = \prod_{r=1}^{\infty} \frac{1}{1-z^r},$$

$$\sum_{n=0}^{\infty} q(n)z^n = \prod_{r=1}^{\infty} (1+z^r) = \prod_{r=1}^{\infty} \frac{1}{1-z^{2r-1}},$$

based on the transformation theory of the elliptic modular functions. Later researches have tended in the direction of a still deeper study of particular problems, culminating in the exact formulae of Rademacher [6] and his followers, or in the direction of a broader and more elementary treatment giving less precise results than (1). But a specification of properties of the generating functions sufficient for the deduction of (1), but not highly extravagant, seems to be lacking.

In this paper we show how to deduce (1) from an elementary knowledge of the asymptotic behaviour of the generating functions when $z(= x + iy)$ approaches the principal singularity $z = 1$ in an arbitrarily wide 'Stolz angle' $|y| \leqq \Delta(1-x)$, $(0 < \Delta < \infty)$. A general result to which the method naturally leads is set out as Theorem 2 in §3, with indications of some of the more immediate applications.

Theorem 2 is deduced from Theorem 1, a Tauberian theorem for the integral

$$(2) \qquad f(s) = \int_0^{\infty} e^{-us}\, dA(u), \qquad A(0) = 0, \qquad (s = \sigma + it),$$

which occupies an intermediate position between theorems of Hardy-Littlewood type with conclusion $A(u) \sim u^a$ and of Wiener-Ikehara type with conclusion $A(u) \sim e^{au}$, $(u \to \infty\,;\, a > 0)$. The nature of the interpolation is exhibited in

the following table, in which a simple specimen of each type of theorem has been selected.

HYPOTHESES		CONCLUSION
(A_1) $f(s) \sim 1/s$ when $s \to 0$ by real positive values;	$A(u)$ increasing;	$A(u) \sim u.$
(A_2) $f(s) \sim \left(\frac{2\pi s}{1-\alpha}\right)^{\frac{1}{2}} \exp\left\{\frac{1-\alpha}{\alpha}\left(\frac{1}{s}\right)^{\alpha/(1-\alpha)}\right\}$ when $s \to 0$ in every fixed 'Stolz angle' $\lvert t \rvert \leqq \Delta\sigma$, $(0 < \Delta < \infty)$;		$A(u) \sim u^{-\frac{1}{2}} e^{u^{\alpha/\alpha}}$, $(0 < \alpha < 1).$
(A_3) $f(s) - 1/(s-1) \to$ limit when $\sigma \to 1+0$, uniformly in every fixed interval $\lvert t \rvert \leqq T$, $(0 < T < \infty)$;		$A(u) \sim e^u.$

Theorem 1, of which (A_2) is a very special case, is stated and proved, on the lines of the Heilbronn-Landau version of the Wiener-Ikehara theorem, in §2.

It will be seen that, in the method of interpolation shown in the above table, the region involved in the hypothesis on $f(s)$ undergoes a discontinuous change of form at each end of the range but no steady change in the interior. This is in contrast to some other possible methods, of which two may be mentioned here. In one method, essentially that of Wiener and Pitt [8], we first transform (A_1) and (A_3) into equivalent theorems (A_1^*) and (A_3^*) by changing '$f(s) - (s-1)^{-1} \to$ limit when $\sigma \to 1+0, \cdots$,' to '$f(s) - s^{-1} \to$ limit when $\sigma \to +0, \cdots$,' in the hypotheses of (A_3) and replacing the conclusions by

$$\frac{A(u+\delta h_u) - A(u-\delta h_u)}{2\delta h_u} \to 1 \quad \text{as} \quad u \to \infty \quad \text{for fixed } \delta, \qquad (0 < \delta \leqq 1),$$

where h_u is u in (A_1^*) and 1 in (A_3^*);[1] and we then interpolate by taking intermediate forms of h_u, say $h_u = u^\vartheta$, $(1 > \vartheta > 0)$. Another method, outlined by Avakumović [1], is to interpolate similarly (by way of h_u) between the convergence theorems associated with (A_1) and (A_3), in which the conclusion is $A(u) \to A$ and the hypothesis on $A(u)$, taken in the 'Schmidt' form, is

$$\lim_{\substack{u, u' \to \infty \\ u \leqq u' \leqq u+\delta h_u}} \{A(u) - A(u')\} \geqq -\epsilon(\delta), \qquad \lim_{\delta \to +0} \epsilon(\delta) = 0,$$

with h_u as before. In these two methods the hypotheses on $f(s)$ (which need not be specified in detail here) involve a region of the form $\lvert t \rvert \leqq C\sigma^\vartheta$ which gradually extends (in the neighbourhood of $s = 0$) as ϑ decreases. This extension of the region is a natural way of tightening the bond between $f(s)$ and its approximation ($1/s$ or A) to balance the strengthening of the conclusion, or the weakening of the hypothesis on $A(u)$, as the order of h_u decreases. But the fact

[1] The equivalence of (A_1) and (A_1^*) is trivial, and that of (A_3) and (A_3^*) is easily proved by setting up the relation $f(s) = f^*(s-1)$, where $f(s)$ and $f^*(s)$ satisfy the conditions of (A_3) and (A_3^*), respectively.

that the boundary of the region touches the imaginary axis at $s = 0$ makes this kind of hypothesis more difficult to verify in applications than the 'Stolz angle' hypotheses used in this paper.

An explicit deduction of the asymptotic formula for $p(n)$ from a Tauberian theorem has been indicated by Avakumović [1, 2], but his demands on the generating function are much more severe, as regards both the region and the degree of the approximation, than those made by Theorem 1.[2]

There is also a paper by Martin and Wiener [5] which works in broadly the same range as Theorem 1 with regard to the rate of growth of $f(s)$ with $1/s$, but on a real variable basis. A passage to the complex domain is effected by means of Parseval's theorem in the case of power series, and the possibility of arithmetical applications is announced. This paper, and a paper by Avakumović and Karamata [3], will be referred to again in §4 in the course of some comments on the conditions of Theorem 1.

I am indebted to Professor Wiener for his help in elucidating the relationship between various methods and results.

NOTATION. We write $f \prec g, f \succ g, f \preccurlyeq g$ for $f/g \to 0, f/g \to \infty, f = O(g)$ respectively, where f and g are positive functions.

A sloping arrow indicates monotonic approach to a (finite or infinite) limit.

We denote the open and closed segments with end points a and b in the complex plane by (a, b) and $[a, b]$, respectively, and use the mixed notations $(a, b]$ and $[a, b)$ with the obvious meanings.

Limit operations in the complex plane are, of course, to be understood in the usual two-dimensional sense. Thus the hypothesis (i) of Theorem 1 (below) means: Given $\epsilon > 0$, we can find $\delta = \delta(\epsilon)$ so that $|f(s)/f_0(s) - 1| < \epsilon$ at all points s of D for which $0 < |s| < \delta$.

2. A Tauberian theorem

Let there be given two functions $\varphi(s)$ and $\chi(s)$, and a domain D of the s-plane, satisfying the following conditions:

a) *$\varphi(s)$ and $\chi(s)$ are regular in D, and real and positive on a segment $(0, h]$ of the positive real axis lying in D;*

b) $-\sigma\varphi'(\sigma) \nearrow \infty$ as $\sigma \searrow 0$

[whence $\{\sigma\varphi'(\sigma)\}' \geqq 0$ and therefore

$$\sigma\varphi''(\sigma) \geqq -\varphi'(\sigma) > 0 \tag{3}$$

for sufficiently small positive σ];

c) $$\frac{\delta(\sigma)}{\sigma} \succ \frac{\{\varphi''(\sigma)\}^{\frac{1}{2}}}{-\varphi'(\sigma)} \quad \text{as} \quad \sigma \searrow 0,$$

[2] In a more recent paper, *Neuer Beweis eines Satzes von G. Hardy und S. Ramanujan über das asymptotische Verhalten der Zerfällungskoeffizienten*, Am. Jour. of Math., **62** (1940), pp. 877–880, Avakumović applies substantially the same methods directly to $p(n)$ without explicit formulation of the relevant Tauberian theorem.

where $\delta(\sigma)$ is the distance of the point σ from the complement of D [so that

$$\delta(\sigma) \leqq \sigma, \qquad (0 < \sigma \leqq h), \tag{4}$$

since the origin cannot belong to D by a) and b)];

d) $$\varphi''(\sigma + z) = O\{\varphi''(\sigma)\},$$
$$\chi(\sigma + z) = O\{\chi(\sigma)\},$$

uniformly for $|z| < \delta(\sigma)$ when $\sigma \searrow 0$.

Define:

$$f_0(s) = \chi(s)e^{\varphi(s)}, \qquad F_0(s) = f_0(s)/s,$$
$$A_0(\omega) = \frac{\chi(\sigma)e^{\varphi(\sigma)+\omega\sigma}}{\{2\pi\sigma^2\varphi''(\sigma)\}^{\frac{1}{2}}} = \frac{F_0(\sigma)e^{\omega\sigma}}{\{2\pi\varphi''(\sigma)\}^{\frac{1}{2}}},$$

where $\sigma = \sigma_\omega$ is the solution [existent and unique for sufficiently large ω, by a) and b)] of

$$-\varphi'(\sigma) = \omega, \qquad (0 < \sigma \leqq h). \tag{5}$$

In these circumstances we have

THEOREM 1. *Suppose that the integral (2) is convergent for $\sigma > 0$, and that*

(i) *$f(s) \sim f_0(s)$ when $s \to 0$ in D;*

(ii) *$f(s) = O\{f_0(|s|)\}$ when $s \to 0$ in some fixed angle 'Δ' of the form $|t| \leqq \Delta\sigma$, $(0 < \Delta < \infty)$;*

(iii) *$A(u)$ is increasing (in the wide sense) for $u \geqq 0$.*

Then

$$\mathfrak{p}(\Delta) \leqq \varlimsup_{u\to\infty} \frac{A(u)}{A_0(u)} \leqq \mathfrak{P}(\Delta), \tag{6}$$

where $\mathfrak{p}(\Delta)$ and $\mathfrak{P}(\Delta)$ depend only on Δ, are strictly positive and finite for any given Δ, $(0 < \Delta < \infty)$, and tend to 1 when $\Delta \to \infty$.

If (ii) *holds for every fixed Δ, then*

$$A(u) \sim A_0(u) \quad as \quad u \to \infty. \tag{7}$$

The last part is obtained from (6) by taking Δ arbitrarily large, and we therefore concentrate on (6). We begin with two lemmas concerning $f_0(s)$.

LEMMA 1. *We have, uniformly for $|z| \leqq \tau(\sigma) < \delta(\sigma)$, when $\sigma \searrow 0$,*

$$\frac{f_0(\sigma + z)}{f_0(\sigma)} = e^{z\varphi'(\sigma)+\frac{1}{2}z^2\varphi''(\sigma)\{1+o(1)\}+o(1)}$$
$$= e^{z\varphi'(\sigma)+\frac{1}{2}z^2\varphi''(\sigma)}\{1 + o(|z|^2\varphi''(\sigma) + 1)e^{\epsilon|z|^2\varphi''(\sigma)}\}$$

for any fixed $\epsilon > 0$.

By the definition of $f_0(s)$,

$$\frac{f_0(\sigma + z)}{f_0(\sigma)} = \frac{\chi(\sigma + z)}{\chi(\sigma)}\, e^{\varphi(\sigma+z)-\varphi(\sigma)}. \tag{8}$$

By a) and d) we have

$$\chi'(\sigma+z)=\frac{1}{2\pi i}\int_\Gamma \frac{\chi(\sigma+\zeta)}{(\zeta-z)^2}\,d\zeta = O\left(\frac{\chi(\sigma)}{\delta(\sigma)}\right), \tag{9}$$

where Γ is the circle $|\zeta - z| = \frac{1}{2}\delta(\sigma)$, and all estimates are uniform for $|z| \leqq \tau(\sigma) < \delta(\sigma)$ when $\sigma \searrow 0$; whence by integration

$$\chi(\sigma+z)-\chi(\sigma)=O\left(\frac{\chi(\sigma)\,|z|}{\delta(\sigma)}\right)=O\left(\frac{\chi(\sigma)\tau(\sigma)}{\delta(\sigma)}\right)=o\{\chi(\sigma)\}. \tag{10}$$

Similarly $\varphi''(\sigma+z)-\varphi''(\sigma)=o\{\varphi''(\sigma)\}$, whence by two integrations

$$\varphi(\sigma+z)-\varphi(\sigma)=z\varphi'(\sigma)+\tfrac{1}{2}z^2\varphi''(\sigma)+o\{|z|^2\varphi''(\sigma)\}. \tag{11}$$

Substituting from (10) and (11) into (8) we obtain the first formula of the lemma, and the second follows from it in virtue of the inequality $|e^Z - 1| \leqq |Z|\,e^{|Z|}$.

LEMMA 2. *$f_0(\sigma)$ increases as σ decreases, for sufficiently small positive σ.*

Applying (9) with $z = 0$, we obtain

$$\frac{-f_0'(\sigma)}{f_0(\sigma)}=-\varphi'(\sigma)-\frac{\chi'(\sigma)}{\chi(\sigma)}=-\varphi'(\sigma)+O(1/\delta(\sigma))\sim-\varphi'(\sigma) \tag{12}$$

as $\sigma \searrow 0$, since, by c), (3) and b),

$$-\delta(\sigma)\varphi'(\sigma) > \{\sigma^2\varphi''(\sigma)\}^{\frac{1}{2}} \geqq \{-\sigma\varphi'(\sigma)\}^{\frac{1}{2}} > 1.$$

Thus $-f_0'(\sigma) > 0$ ultimately. This proves the lemma, but we carry the argument a little further for future reference.

Given ϵ, $(0 < \epsilon < 1)$, we have, by (12) and b),

$$\frac{1}{\epsilon\sigma} < -(1-\epsilon)\varphi'(\sigma) < -\frac{f_0'(\sigma)}{f_0(\sigma)} < -(1+\epsilon)\varphi'(\sigma), \qquad (0 < \sigma \leqq \sigma_\epsilon),$$

whence, integrating over $[\sigma, \sigma_\epsilon]$ and taking exponentials,

$$A_\epsilon\sigma^{-1/\epsilon} < B_\epsilon e^{(1-\epsilon)\varphi(\sigma)} < f_0(\sigma) < C_\epsilon e^{(1+\epsilon)\varphi(\sigma)}, \qquad (0 < \sigma < \sigma_\epsilon), \tag{13}$$

where $A_\epsilon, B_\epsilon, C_\epsilon > 0$.

PROOF OF THEOREM 1. In the following argument the main variables are σ and ω, connected by (5) and supposed 'sufficiently small' and 'sufficiently large', respectively. Limit operations refer to the process $\sigma \searrow 0$ or the equivalent process $\omega \nearrow \infty$.

In the formula

$$\int_0^\infty A(u)e^{-us}\,du=\frac{f(s)}{s}=F(s), \qquad (\sigma > 0),$$

say, obtained from (2) by partial integration, change s to $\sigma + z$, multiply by $k(z/T)e^{\xi z}\,dz$, where $T > 0$, ξ is real, and $k(z)$ is any function regular along $[0, i]$, and integrate along a path l from 0 to iT. This gives

$$\int_0^\infty A(u)e^{-u\sigma}K(Tu-T\xi)T\,du=\int_0^{iT}F(\sigma+z)k(z/T)e^{\xi z}\,dz, \tag{14}$$

where

$$K(v) = \int_0^i k(Z)e^{-vZ}\,dZ \tag{15}$$

taken along the path L corresponding to l by the transformation $z = TZ$, provided that $\Re(\sigma + z) > 0$ at all points z of l. For L we may take the segment $[0, i]$ or any sufficiently near path giving the same value of $K(v)$.

Take

$$\xi = \xi(\sigma) = \omega + O(\sigma^{-1}) = -\varphi'(\sigma) + O(\sigma^{-1}), \tag{16}$$

$$T = T(\sigma) = \Delta\sigma, \tag{17}$$

with the Δ of hypothesis (ii), and suppose $k(0) \neq 0$. Let $l = l(\sigma)$ be made up of the segment l_1 from 0 to $i\tau$ and the broken line l_2 from $i\tau$ to iT formed of two segments to the left of, and inclined at $\tan^{-1}\tau/\sigma$ to, the segment $[i\tau, iT]$, where $\tau = \tau(\sigma)$ is chosen so that

$$\delta(\sigma) > \tau(\sigma) > \frac{\sigma\{\varphi''(\sigma)\}^{\frac{1}{2}}}{-\varphi'(\sigma)}. \tag{18}$$

This choice is possible by c); and the path $L = L(\sigma)$ corresponding to $l = l(\sigma)$ tends to coincidence with the segment $[0, i]$, its greatest distance from this segment being $O(\tau/\sigma) = o(1)$ by (18) and (4).

When $z = iy$ is on l_1, the point $\sigma + z$ is in D, by (18), whence, by (16), (17) and hypothesis (i) [since $|z/\sigma| \leqq \tau/\sigma = o(1)$],

$$\frac{F(\sigma+z)k(z/T)e^{\xi z}}{F_0(\sigma)k(0)} = \frac{f(\sigma+z)}{f_0(\sigma)}\cdot\frac{\sigma k(z/\Delta\sigma)}{(\sigma+z)k(0)}\cdot e^{\omega z + O(\tau/\sigma)}$$

$$= \frac{f_0(\sigma+z)}{f_0(\sigma)}\,e^{o(1)}\cdot e^{o(1)}\cdot e^{-z\varphi'(\sigma)+o(1)} = e^{-\frac{1}{2}y^2\varphi''(\sigma)}\{1 + o(y^2\varphi''(\sigma)+1)e^{\frac{1}{2}y^2\varphi''(\sigma)}\}$$

uniformly in y, by Lemma 1 (with $\epsilon = \frac{1}{4}$). Inserting this in (14) and making the substitution $y\{\varphi''(\sigma)\}^{\frac{1}{2}} = w$, we deduce that the contribution of l_1 to the right hand side of (14) is

$$\frac{ik(0)F_0(\sigma)}{\{\varphi''(\sigma)\}^{\frac{1}{2}}}\left(\int_0^{\eta(\sigma)} e^{-\frac{1}{2}w^2}\,dw + o\int_0^{\eta(\sigma)}(w^2+1)e^{-\frac{1}{2}w^2}\,dw\right) \sim \frac{ik(0)F_0(\sigma)(\pi/2)^{\frac{1}{2}}}{\{\varphi''(\sigma)\}^{\frac{1}{2}}},$$

where $\eta(\sigma) = \tau(\sigma)\{\varphi''(\sigma)\}^{\frac{1}{2}} > 1$ by (18) and (3).

When $z = x + iy$ is on l_2, the point $\sigma + z$ is in the angle 'Δ' of hypothesis (ii) and $|\sigma + z| > \sigma$ (as may be seen at once from a figure, the lower segment of l_2 being at right angles to the segment $[-\sigma, i\tau]$); whence, by (ii), (16), (17) and Lemma 2,

$$F(\sigma+z)k(z/T)e^{\xi z} = O\left(\frac{f_0(|\sigma+z|)}{|\sigma+z|}\cdot 1\cdot e^{\omega x + O(T/\sigma)}\right) = O\{F_0(\sigma)e^{\omega x}\}$$

uniformly in z. Also $x \leqq 0$ and

$$\left|\frac{dz}{dx}\right| = |1 \mp (\sigma/\tau)i| = O(\sigma/\tau)$$

on each of the segments of l_2. Hence the contribution of l_2 to the right hand side of (14) is

$$O\left(F_0(\sigma)\int_{-\infty}^{0} e^{\omega x}(\sigma/\tau)\,dx\right) = O\left(\frac{F_0(\sigma)\sigma}{\omega\tau}\right) = o\left(\frac{F_0(\sigma)}{\{\varphi''(\sigma)\}^{\frac{1}{2}}}\right),$$

since $\omega\tau = -\varphi'(\sigma)\tau(\sigma) \succ \sigma\{\varphi''(\sigma)\}^{\frac{1}{2}}$ by (18). Thus

$$\int_0^{iT} F(\sigma+z)k(z/T)e^{\xi z}\,dz \sim \frac{ik(0)(\pi/2)^{\frac{1}{2}}F_0(\sigma)}{\{\varphi''(\sigma)\}^{\frac{1}{2}}} = ik(0)\pi A_0(\omega)e^{-\omega\sigma},$$

by the definition of $A_0(\omega)$. Substituting in (14) and putting $u = \xi + v/T$, we obtain

$$\int_{-T\xi}^{\infty} A(\xi + v/T)e^{-v\sigma/T}K(v)\,dv \sim ik(0)\pi A_0(\omega)e^{(\xi-\omega)\sigma}, \tag{19}$$

in which we note that, by (16), (17) and b),

$$T\xi = -\Delta\sigma\varphi'(\sigma) + O(1) \succ 1. \tag{20}$$

Take, firstly

$$k(Z) = 2(Z - i).$$

Then

$$K(v) = -\frac{2i}{v} + \frac{2(1-e^{-iv})}{v^2}, \qquad \Re K(v) = \left(\frac{\sin v/2}{v/2}\right)^2 \geqq 0.$$

Taking real parts in (19), and using (20) and hypothesis (iii) [which implies $A(u) \geqq A(0) = 0$, $(u \geqq 0)$], we deduce that

$$A(\xi - \lambda/T)e^{-\lambda\sigma/T}\int_{-\lambda}^{\lambda}\left(\frac{\sin v/2}{v/2}\right)^2 dv < \{1 + o(1)\}2\pi A_0(\omega)e^{(\xi-\omega)\sigma},$$

where λ is any fixed positive number. Taking

$$\xi = \omega + \lambda/T = \omega + \lambda/\Delta\sigma$$

[an admissible choice by (16)], we infer that

$$\overline{\lim}\,\frac{A(\omega)}{A_0(\omega)} \leqq \frac{2\pi e^{2\lambda/\Delta}}{\int_{-\lambda}^{\lambda}\left(\frac{\sin v/2}{v/2}\right)^2 dv},$$

and the second part of (6) follows on putting $\lambda = \Delta^{\frac{1}{2}}$.

Take, secondly,

$$k(Z) = 2(Z - i) - \frac{e^{\mu Z} - e^{-\mu Z}}{\mu}, \qquad \mu > 0, \qquad \mu \equiv 0 (\mathrm{mod}\ 2\pi).$$

Then

$$K(v) = -\frac{2i}{v} + (1 - e^{-iv})\left(\frac{2}{v^2} - \frac{1}{\mu(v-\mu)} + \frac{1}{\mu(v+\mu)}\right),$$

$$\Re K(v) = \left(\frac{\sin v/2}{v/2}\right)^2 \frac{\mu^2}{\mu^2 - v^2} \begin{cases} \geqq 0, & (|v| \leqq \mu), \\ \leqq 0, & (|v| \geqq \mu), \end{cases}$$

whence, by a similar argument, with $\xi = \omega - \mu/T$

$$\underline{\lim} \frac{A(\omega)}{A_0(\omega)} \geqq \frac{2\pi e^{-2\mu/\Delta}}{\int_{-\mu}^{\mu} \left(\frac{\sin v/2}{v/2}\right)^2 \frac{\mu^2}{\mu^2 - v^2} dv}.$$

The denominator here is

$$\int_{-\mu}^{\mu} 2 \sin^2 v/2 \left(\frac{2}{v^2} + \frac{1}{\mu(\mu - v)} + \frac{1}{\mu(\mu + v)}\right) dv$$

$$= \int_{-\mu}^{\mu} \left(\frac{\sin v/2}{v/2}\right)^2 dv + \frac{4}{\mu} \int_0^{2\mu} \frac{\sin^2 w/2}{w} dw,$$

and the first part of (6) follows on taking $\mu = 2\pi([\Delta^{\frac{1}{2}}] + 1)$.

An alternative treatment of the range from $i\tau$ to iT in the above proof is possible under certain conditions, for example on the additional assumption b′) of §4 (Theorem 1′), and, in particular, in the special case to be applied in §3. This consists in taking the segment $[i\tau, iT]$ as path of integration and applying an estimate

$$\frac{F(\sigma + iy)k(iy/T)e^{i\xi y}}{F_0(\sigma)} = O(e^{\frac{1}{2}(|\sigma+iy|-\sigma)\phi'(|\sigma+iy|)})$$

$$= O(e^{\delta\sigma^{-1}y^2\phi'(\sigma)}) = O\left(\frac{\sigma}{-\varphi'(\sigma)y^2}\right), \qquad (\delta > 0)$$

to the integrand. [The first step here depends on (ii) and an application of the mean value theorem to $\log f_0(\sigma)$, followed by (12) and b); but the next step seems to require an additional assumption such as b′).] The procedure places the argument on an essentially 'real variable' footing, and this may be considered desirable with a view to the logical status of Theorem 1 in arithmetical applications. It should be observed, however, that the use of a complex variable technique in the original (general) proof does not imply any essential dependence on complex function theory. Thus the fact that the special integrals $K(v)$ are independent of the path $L = L(\sigma)$ is verified by the calculations without reference to Cauchy's general theorem. And, of course, the incidental appeals to Cauchy's theorem in Lemmas 1 and 2 can be avoided in particular applications, like that of §3, by an independent verification of the results.

Special Cases. Take first

$$\varphi(s) = \beta^{-1}(M/s)^{\beta}, \qquad \chi(s) = C(M/s)^{m\beta - \frac{1}{2}}, \qquad (\beta, M, C > 0, m \text{ real}).$$

Then

$$f_0(s) = C(M/s)^{m\beta - \frac{1}{2}} e^{\beta^{-1}(M/s)^{\beta}}, \tag{21}$$

$$A_0(\omega) = \left(\frac{1 - \alpha}{2\pi}\right)^{\frac{1}{2}} C(M\omega)^{m\alpha - \frac{1}{2}} e^{\alpha^{-1}(M\omega)^{\alpha}}, \qquad (\alpha = \beta/(\beta + 1)), \tag{22}$$

as may be verified by a simple calculation. For D we may take, for example, any domain of the form $|t| < \Delta_0\sigma^\gamma$ with $\Delta_0 > 0$, $\gamma \geqq 1$ [to satisfy d)], and $\gamma < 1 + \beta/2$ [to satisfy c)]. The actual result to be used in §3, and quoted with special values of M, m, C as (A_2) in §1, is obtained directly by taking $\gamma = 1$ and Δ_0 arbitrarily large; but the hypotheses (i) and (ii) on $f(s)$ become more general as γ is increased (within the permitted range $1 \leqq \gamma < 1 + \beta/2$). Actually it will appear from §4 (Theorem 1′) that in this example we may replace (i) by the corresponding hypothesis along the positive real axis, retaining the complex variable explicitly only in the hypothesis (ii).

The scope of the conditions a), b), c), d) may be illustrated by two further examples, in which we take $\chi(s) = 1$:

$$\varphi(s) = l_1(1/s)l_n(1/s), \qquad \sigma\{l_n(1/\sigma)\}^{-\frac{1}{2}} < \delta(\sigma) \leqq \vartheta\sigma, \qquad (0 < \vartheta < 1);$$

$$\varphi(s) = e_n(1/s), \qquad \sigma\{e_n(1/\sigma)\}^{-\frac{1}{2}} < \delta(\sigma) \lesssim \sigma^2\{e_1(1/\sigma)\cdots e_{n-1}(1/\sigma)\}^{-1}.$$

Here l_n and e_n denote iterated logarithms and exponentials, and the conditions on $\delta(\sigma)$ correspond to c) and d), respectively. A study of such examples, in conjunction with the inequalities (13), shows that $f_0(\sigma)$ must increase more rapidly than any power of $1/\sigma$ when $\sigma \searrow 0$, but that, subject to conditions of regularity and smoothness, any rate of growth above this limit is permitted, the latitude in the choice of D becoming greater with the rate of growth. The extraneous factor $\chi(s)$ adds nothing essential, since it can in practice be absorbed by adding $\log \chi(s)$ to $\varphi(s)$; it serves merely to simplify the calculation of $A_0(\omega)$ in special cases.

3. Application to partitions

THEOREM 2. *Let*

$$0 < \lambda_1 < \lambda_2 < \cdots$$

be a given sequence with

$$N(u) = Bu^\beta + R(u), \qquad (B > 0, \beta > 0), \tag{23}$$

where $N(u)$ is the number of λ_r not exceeding u, and

$$\int_0^u \frac{R(v)}{v}\,dv = b\log u + c + o(1) \qquad \text{as } u \to \infty. \tag{24}$$

For real l let $p(l)$ be the number of solutions of

$$l = r_1\lambda_1 + r_2\lambda_2 + \cdots$$

in integers $r_r \geqq 0$, and $p^*(l)$ the number of solutions with $r_r = 0$ or 1, (where the formally infinite sum contains a finite but unrestricted number of non-zero terms). For u real and $h > 0$ let

$$P(u) = \sum_{l<u} p(l), \qquad P^*(u) = \sum_{l<u} p^*(l),$$

$$P_h(u) = \frac{P(u) - P(u-h)}{h}, \qquad P_h^*(u) = \frac{P^*(u) - P^*(u-h)}{h}$$

where each summation is over the (discrete) set of l for which the summand is different from zero.

Then, when $u \to \infty$,

$$P(u) \sim \left(\frac{1-\alpha}{2\pi}\right)^{\frac{1}{2}} e^{c} M^{-(b+\frac{1}{2})\alpha} u^{(b+\frac{1}{2})(1-\alpha)-\frac{1}{2}} e^{\alpha^{-1}(Mu)^{\alpha}}, \tag{25}$$

$$P^*(u) \sim \left(\frac{1-\alpha}{2\pi}\right)^{\frac{1}{2}} 2^{b} (M^*u)^{-\alpha/2} e^{\alpha^{-1}(M^*u)^{\alpha}}, \tag{26}$$

where

$$\alpha = \frac{\beta}{\beta+1}, \qquad M = \{B\beta\Gamma(\beta+1)\zeta(\beta+1)\}^{1/\beta}, \qquad M^* = (1-2^{-\beta})^{1/\beta} M.$$

Also

$$P_h(u) \sim \left(\frac{1-\alpha}{2\pi}\right)^{\frac{1}{2}} e^{c} M^{-(b-\frac{1}{2})\alpha} u^{(b-\frac{1}{2})(1-\alpha)-\frac{1}{2}} e^{\alpha^{-1}(Mu)^{\alpha}}, \tag{27}$$

$$P_h^*(u) \sim \left(\frac{1-\alpha}{2\pi}\right)^{\frac{1}{2}} 2^{b} M^{*\alpha/2} u^{\frac{1}{2}\alpha-1} e^{\alpha^{-1}(M^*u)^{\alpha}}, \tag{28}$$

if h is a positive constant for which the left hand side is an increasing function of u. This condition is certainly satisfied for $P_h(u)$ if h belongs to the given sequence $\{\lambda_\nu\}$, and for $P_h^*(u)$, if h, $2h$, 2^2h, $\cdots$ all belong to $\{\lambda_\nu\}$.

The formulae (26) *and* (28) *for $P^*(u)$ and $P_h^*(u)$ remain valid if* (24) *is replaced by the weaker assumption*

$$\int_0^u R(v)\,dv = bu + o(u), \qquad \text{as } u \to \infty \tag{29}$$

[*derivable from* (24) *by partial integration*].

Consider first $P(u)$ and $P_h(u)$. By (23) and (24)

$$N(u) \leqq 2\int_u^{2u} \frac{N(v)}{v}\,dv = 2\left(\int_u^{2u} \frac{Bv^{\beta}}{v}\,dv + L(2u) - L(u)\right) = O(u^{\beta}),$$

where $L(u)$ is the left hand side of (24). This suffices to justify the formal transformations which follow, when $\sigma > 0$. By the definition of $p(l)$,

$$\sum_l p(l)e^{-ls} = \prod_{\nu=1}^{\infty} (1-e^{-\lambda_\nu s})^{-1} = g(s),$$

say. By (23) and (24),

$$\log g(s) = -\sum_{\nu=1}^{\infty} \log(1-e^{-\lambda_\nu s}) = -\int_0^{\infty} \log(1-e^{-us})\,dN(u)$$

$$= \int_0^{\infty} \frac{s}{e^{us}-1} N(u)\,du = \int_0^{\infty} \frac{s}{e^{us}-1} (Bu^{\beta}\,du + u\,dL(u)) \tag{30}$$

$$= \frac{B}{s^{\beta}} \int_{u=0}^{\infty} \frac{(us)^{\beta}\,d(us)}{e^{us}-1} - \int_{u=0}^{\infty} L(u)\,d\left(\frac{us}{e^{us}-1}\right)$$

$$= \frac{B\Gamma(\beta+1)\zeta(\beta+1)}{s^\beta} - \int_0^\infty (b\log u + c)\psi'(us)s\,du - \int_0^\infty S(u)\psi'(us)s\,du, \tag{31}$$

where $\psi(s) = s/(e^s - 1)$, and $S(u)$ is $o(1)$ when $u \to \infty$, $O(|\log u|)$ when $u \searrow 0$, and continuous for $u > 0$. The first integral in (31) is

$$\left[(b\log u + c)\frac{us}{e^{us}-1} - b\log(1-e^{-us})\right]_{u=0}^{\infty} = b\log s - c.$$

And, since $e^{s/2}\psi'(s)$ and $|s|/\sigma$ are bounded in any fixed angle 'Δ' of the type $|t| \leqq \Delta\sigma$, $(0 < \Delta < \infty)$, the second integral is

$$O\left(\int_0^{1/\sqrt{\sigma}} (|\log u| + 1)\sigma\,du\right) + o\left(\int_{1/\sqrt{\sigma}}^\infty e^{-u\sigma/2}\sigma\,du\right) = o(1)$$

when $s \to 0$ in any 'Δ'. Hence

$$\log g(s) = \beta^{-1}(M/s)^\beta - b\log s + c + o(1),$$

$$g(s) \sim e^c s^{-b} e^{\beta^{-1}(M/s)^\beta}, \tag{32}$$

when $s \to 0$ in any fixed 'Δ', where $M^\beta = B\beta\Gamma(\beta+1)\zeta(\beta+1)$.

But

$$\int_0^\infty e^{-us}\,dP(u) = g(s),$$

$$\int_0^\infty e^{-us}\,dP_h(u) = \frac{1}{h}\left(\int_0^\infty e^{-us}\,dP(u) - \int_0^\infty e^{-(u+h)s}\,dP(u)\right) = \frac{1-e^{-hs}}{h}\,g(s), \tag{33}$$

and $(1-e^{-hs})/h \sim s$ when $s \to 0$ in any way. The results (25) and (27) may now be obtained at once by taking $m = (b \pm \frac{1}{2})/\beta$, $C = M^{\frac{1}{2}-m\beta}e^c$, in the first of the special cases of Theorem 1 discussed at the end of §2, since $P(u)$ obviously increases with u and $P_h(u)$ is explicitly assumed to increase.

Finally, the identity (33), which is equivalent to an identity between Dirichlet's series, shows that the step-function $P_h(u)$ increases with u if and only if the series for $(1-e^{-hs})g(s)$ has positive coefficients; and this is certainly so if h belongs to $\{\lambda_\nu\}$, since then

$$(1-e^{-hs})g(s) = \prod_{\lambda_\nu \neq h} (1-e^{-\lambda_\nu s})^{-1}.$$

The formulae for $P^*(u)$ and $P_h^*(u)$ follow similarly from

$$\sum_l p^*(l)e^{-ls} = \prod_{\nu=1}^\infty (1+e^{-\lambda_\nu s}) = g^*(s),$$

$$g^*(s) \sim 2^b e^{\beta^{-1}(M^*/s)^\beta}, \qquad \text{when } s \to 0 \text{ in any '}\Delta\text{'},$$

where $M^{*\beta} = (1-2^{-\beta})M^\beta$. The last relation may be deduced from (32), since $g^*(s) = g(s)/g(2s)$. But the direct proof is more elementary and demands only (29) in place of (24), because $\log(1+e^{-s})$ and $1/(e^s+1)$ are regular at $s = 0$,

so that in the formula corresponding to (30) we can use $R(u)\,du$ directly without replacing it by $u\,dL(u)$. The step-function $P_h^*(u)$ increases with u if and only if the Dirichlet's series for $(1 - e^{-hs})g^*(s)$ has positive coefficients; and this is certainly so if the numbers $2^\mu h$, $(\mu = 0, 1, 2, \cdots)$, all belong to $\{\lambda_\nu\}$, since then

$$(1 - e^{-hs})g^*(s) = (1 - e^{-hs}) \prod_{\mu=0}^{\infty} (1 + e^{-2^\mu hs}) \prod_{\lambda_\nu \neq 2^\mu h} (1 + e^{-\lambda_\nu s}) = \prod_{\lambda_\nu \neq 2^\mu h} (1 + e^{-\lambda_\nu s}).$$

Special Cases. The formulae (1) may be deduced at once form (27) and (28), respectively, by taking $\lambda_\nu = \nu$, $(\nu = 1, 2, \cdots)$ and $h = 1$, since

$$\int_0^u \frac{[v] - v}{v}\,dv = \sum_{\nu \leq u} \log \frac{u}{\nu} - u = -\tfrac{1}{2} \log u - \tfrac{1}{2} \log 2\pi + o(1) \tag{34}$$

as $u \to \infty$ (by Stirling's theorem for $\log n!$), and $\Gamma(2)\zeta(2) = \pi^2/6$. In the case of $q(n)$ the simpler result

$$\int_0^u ([v] - v)\,dv = -\tfrac{1}{2}u + o(u)$$

may be used in place of (34), in virtue of the concluding remark of Theorem 2. The formula for $q(n)$ may, alternatively, be deduced from (27) by taking $\lambda_\nu = 2\nu - 1$, $h = 1$.

An asymptotic formula for the number of partitions of a positive integer n into positive integral k-th powers, where k is any given positive integer, may be deduced similarly from (27) by taking $\lambda_\nu = \nu^k$ and $h = 1$, the verification of (23) and (24), with $\beta = 1/k$, $B = 1$, $b = -\frac{1}{2}$, $c = -\frac{1}{2}k \log 2\pi$ depending on a formula equivalent to (34). This particular problem has been studied in detail by Wright [9], who obtains by more special methods an asymptotic expansion in functions of descending order of magnitude.

Some obvious extensions of the case $\lambda_\nu = \nu^k$ fall within the scope of Theorem 2. Thus we may restrict ν to certain arithmetical progressions, and modify a finite portion of the sequence in any way. In applications of (27) or (28) attention must, of course, be paid to the condition that the left hand side increases with u, at any rate for large u (this being obviously sufficient).

Theorem 2 is a very special deduction from Theorem 1 and could easily be extended to cover other partition functions, such as the number of partitions of n into k_1-th and k_2-th powers.

4. Notes on the conditions of Theorem 1

1) Theorem 1′. *If, in addition to the conditions* a), b), c), d), *we have*

b′) $-\sigma^k \varphi'(\sigma) \searrow 0$ *as* $\sigma \searrow 0$, *for some fixed* k, *then the hypothesis* (i) *of Theorem* 1 *may be replaced by*

(i′) $f(s) \sim f_0(s)$ *when* $s \to 0$ *by real positive values.*

We shall prove the following facts:

(α) On the assumptions a), b), c), d) only, (i)′ and (iii) [or, alternatively,

(i)′ and (ii)] imply (i) with D replaced by a certain sub-domain D_f (depending on f) for which

$$\delta_f(\sigma) > \{\varphi''(\sigma)\}^{-\frac{1}{2}}, \qquad \text{as } \sigma \searrow 0, \tag{35}$$

where $\delta_f(\sigma)$ has the same meaning for D_f as $\delta(\sigma)$ for D;

(β) On the additional assumption b′), the domain D_f is 'admissible' in the sense of conditions a), c), d).

An appeal to Theorem 1 with D replaced by D_f will then establish Theorem 1′.

Choose $\rho(\sigma)$ so that (when $\sigma \searrow 0$)

$$\{\varphi''(\sigma)\}^{-\frac{1}{2}} < \rho(\sigma) < \delta(\sigma), \tag{36}$$

as is possible by c) and (3). Consider

$$G(z) = G_\sigma(z) = \frac{f(\sigma + z)}{f_0(\sigma + z)} - 1, \qquad (z = x + iy)$$

in the circle $C(\sigma)$ defined by $|z| \leqq \rho(\sigma)$.

By (iii) and (i)′, $|f(\sigma + z)| \leqq f(\sigma + x) = O\{f_0(\sigma + x)\}$, [or, alternatively, by (ii) and Lemma 2, $f(\sigma + z) = O\{f_0(|\sigma + z|)\} = O(f_0(\sigma + x))$], uniformly for $|z| \leqq \rho(\sigma)$ when $\sigma \searrow 0$; whence, by Lemma 1, $G(z)$ is regular and satisfies

$$|G(z)| = O\left(\frac{f_0(\sigma + x)}{|f_0(\sigma + z)|}\right) + O(1)$$

$$= O(e^{x\varphi'(\sigma)+\frac{1}{2}x^2\varphi''(\sigma)}\, e^{-\Re z\varphi'(\sigma)+\frac{1}{2}|z|^2\varphi''(\sigma)}) + O(1) < e^{2\rho^2(\sigma)\varphi''(\sigma)}$$

in $C(\sigma)$ (if σ is sufficiently small).

By (i)′ we have, since $\rho(\sigma) < \frac{1}{2}\delta(\sigma) \leqq \frac{1}{2}\sigma$ ultimately,

$$|G(z)| < \epsilon(\sigma)$$

on the diameter $[-\rho(\sigma), \rho(\sigma)]$, where $\epsilon(\sigma) = \epsilon_f(\sigma) \to 0$ as $\sigma \searrow 0$ and is independent of the choice of $\rho(\sigma)$ in the range (36).

It follows, by a well known convexity theorem [applied separately to the upper and lower halves of $C(\sigma)$], that, in the region bounded by the two circular arcs $-\rho(\sigma)$, $\pm i\lambda\rho(\sigma)$, $\rho(\sigma)$, $(0 < \lambda < 1)$, and so in particular in the circle $|z| < \lambda\rho(\sigma)$,

$$|G(z)| < \epsilon^{\vartheta}(\sigma) e^{(1-\vartheta)2\rho^2(\sigma)\varphi''(\sigma)}, \tag{37}$$

where $(1 - \vartheta)\pi$ is the angle at which the two circular arcs intersect, so that $0 < \vartheta < 1$ and ϑ depends only on λ.

Take a fixed λ and choose $\rho(\sigma) = \rho_f(\sigma)$ in the range (36) so that the right hand side of (37) is $< \epsilon^{\frac{1}{2}\vartheta}(\sigma)$. Then we shall have $f(s) \sim f_0(s)$ when $s \to 0$ in the sub-domain

$$D_f = \sum_{0<\sigma\leqq h_f} d_f(\sigma)$$

of D, where $d_f(\sigma)$ is the circular domain of the s-plane with centre σ and radius

$\lambda\rho_f(\sigma)$, h_f is a sufficiently small positive number, and $\sum$ denotes logical summation over σ. Also $\delta_f(\sigma) \geqq \lambda\rho_f(\sigma) > \{\varphi''(\sigma)\}^{-\frac{1}{2}}$ by (36). This proves (α).

Since D_f lies in D, contains the segment $(0, h_f]$, and satisfies (35), it will satisfy all the conditions imposed on D in a), c), d), if $\sigma\varphi''(\sigma) = O\{-\varphi'(\sigma)\}$. But this is so on the assumption b'), which implies that $\{\sigma^k\varphi'(\sigma)\}' \leqq 0$ and therefore

$$\sigma\varphi''(\sigma) \leqq -k\varphi'(\sigma).$$

This proves (β) and completes the proof of Theorem 1'.

2) In Theorem 1' we may regard (iii) and (ii) (for every Δ) as Tauberian conditions which convert the generally false inference (i)' $\rightarrow$ (7) into a true proposition. An example showing that (iii) alone is not a sufficient Tauberian condition has been constructed by Avakumovic and Karamata [3] (353, e). In this example (i)' and (iii) are satisfied but the $\underline{\lim}$ and $\overline{\lim}$ in (6) are 0 and ∞.

Another example proving less in this direction, but illustrating the effect of the supplementary hypothesis (ii), may be constructed as follows. With the special functions (21) and (22), and any fixed $U > 0$, we can show, by the saddle point method, that

$$\int_U^\infty e^{-us}\,dA_0(u) \sim f_0(s), \tag{38}$$

when $s \rightarrow 0$ in a certain angle $|t| \leqq \delta\sigma$, $(\delta > 0)$; whence, by a change of the variables and of the path of integration,

$$\int_U^\infty e^{-us}\,dA_0(\rho u) \sim f_0(s/\rho), \tag{39}$$

when $s \rightarrow 0$ in $|t| \leqq \delta'\sigma$, $(0 < \delta' < \delta)$, if $\rho = re^{i\vartheta}$ is fixed with $r > 0$, and $\vartheta > 0$ and sufficiently small.

Choose ρ so that $\Re\rho^\alpha = 1$. Then $\Re\rho^\beta < 1$, since

$$(\cos\beta\vartheta)^{1/\beta} < (\cos\alpha\vartheta)^{1/\alpha},$$

as may be verified without difficulty (for small positive ϑ) since $\beta > \alpha > 0$. Take

$$A(u) = A_0(u) + \lambda\Re A_0(\rho u), \quad (u \geqq U), \qquad A(u) = 0, \quad (0 \leqq u < U),$$

with $\lambda > 0$. Then, since $\Re\rho^\alpha = 1$,

$$\overline{\lim_{u\rightarrow\infty}} \frac{A(u)}{A_0(u)} = 1 \pm \lambda', \qquad (\lambda' > 0). \tag{40}$$

so that (7) is false. On the other hand, since $\Re\rho^\beta < 1$, $f_0(\sigma/\rho) = o\{f_0(\sigma)\}$ as $\sigma \searrow 0$, whence, by (38) and (39),

$$f(s) = \int_0^\infty e^{-us}\,dA(u) \sim f_0(s) \tag{41}$$

when $s \to 0$ by real positive values, so that (i)′ is satisfied. Also, (iii) is satisfied if λ is sufficiently small and U sufficiently large, as may be verified by differentiation.

In this example, (41) holds in the angle $|t| \leqq \Delta\sigma$ if Δ is sufficiently small, so that we have (ii) and therefore (6) for some $\Delta > 0$. On the other hand, (40) shows that, apart from the precise forms of $\mathfrak{p}(\Delta)$ and $\mathfrak{P}(\Delta)$, the conclusion (6) cannot be improved, at any rate for the smaller values of Δ.

The actual form of the supplementary hypothesis (ii) was adopted with a view to applications, but it is naturally not the only possible form.

3) The distinction between Theorem 1′ and the Hardy-Littlewood case $f_0(s) = s^{-a}$, $(a > 0)$, in which (iii) alone is a sufficient Tauberian condition for an inference of the form (i)′ $\to$ (7), can be explained very simply in terms of the ideas of this paper. A valid (if somewhat sophisticated) proof of the Hardy-Littlewood theorem [quoted in the case $a = 1$ as (A_1) in §1] can be constructed by applying, first, a convexity argument, substantially as in Theorem 1′ (α) [with $\chi(s) = 1$ and $\varphi(s) = a \log 1/s$] or as a Phragmén-Lindelöf theorem, to deduce from (i)′ and (iii) that $f(s) \sim f_0(s)$ when $s \to 0$ in any fixed angle $|t| \leqq \Delta\sigma$, $(0 < \Delta < \infty)$, and, then, a slightly modified form of the argument of Theorem 1, with $\tau = T = \Delta\sigma$ and an arbitrarily large Δ, to deduce (7) with $A_0(u) = u^a/\Gamma(a+1)$. Thus the convexity argument is completely successful in removing the complex variable from the hypotheses in the Hardy-Littlewood case, but only partially successful in Theorem 1′, the essential difference lying in the fact that $\{\varphi''(\sigma)\}^{-\frac{1}{2}}$ is of the same order as σ in the former case, but of lower order in the latter.

4) If we ask what can be proved about $A(u)$ when we simply omit (ii) from the hypotheses of Theorem 1′, or, more generally, when we replace the hypotheses (i) and (ii) of Theorem 1 by (i)′, a partial answer can be given by the methods of this paper, though the question belongs more naturally to the order of ideas of Martin and Wiener [5]. We can first use the convexity argument of Theorem 1′(α) to extend (i)′ to the region $|t| \leqq C\{\varphi''(\sigma)\}^{-\frac{1}{2}}$, where C is any positive constant, and then take $\tau = T = C\{\varphi''(\sigma)\}^{-\frac{1}{2}}$ in the argument of Theorem 1 [or in the corresponding argument with $A(u)\,du$ replaced by $dA(u)$ and $F(s)$ by $f(s)$]. The general effect of this, when we reject the parts of the range of integration where $\Re K(v)$ is relatively small and use the condition $A(u) \geqq 0$ [or the condition '$dA(u) \geqq 0$'], is to give information about certain weighted averages of $A(u)\,du$ [or of $dA(u)$] over an interval about the point $u = \omega = -\varphi'(\sigma)$ whose length is of order $1/T$, i.e. of order $\{\varphi''(\sigma)\}^{\frac{1}{2}}$. But this is just the kind of result that Martin and Wiener obtain by real variable methods more appropriate to the form of the problem. The actual theorems proved by Martin and Wiener do not emerge readily by the method outlined above, which is somewhat artificial in this context, but the discussion serves at any rate to explain the effect of the various hypotheses concerning $f(s)$. The extension from the real form (i)′ to the complex form '(i) + (ii)' reduces the length of the interval of averaging from $c\{\varphi''(\sigma)\}^{\frac{1}{2}}$ to c/σ, and this enables us to draw the conclusion (6) when $A(u)$

is monotonic, because $e^{-u\sigma}$ varies only by a constant factor $e^{\pm c}$, when u ranges over an interval of length c/σ; and the sharper conclusion (7) is possible if c can be taken arbitrarily small, because $e^{\pm c} \to 1$ when $c \to 0$.

THE INSTITUTE FOR ADVANCED STUDY
PRINCETON, NEW JERSEY

REFERENCES

(1) V. G. Avakumović, *Über das Verhalten Laplacescher Integrale an der Konvergenzgrenze mit neuem Beweis eines Satzes von Hardy-Ramanujan über das asymptotische Verhalten der Zerfällungskoeffizienten*, (2 Congr. Interbalkan, des Math., Bucarest, 12-IX-1937), Bull. Math. Soc. Roum. Sci. **40**, Nr. 1/2 (1938), 101–106.

(2) V. G. Avakumović, *Bemerkungen über Laplacesche Integrale, deren Wachstum von Exponentialcharakter ist.* I *und* II, Math. Zeitschrift, **46** (1940), 62–69.

(3) V. Avakumović and J. Karamata, *Über einige Taubersche Sätze, deren Asymptotik von Exponentialcharakter ist.* I, Math. Zeitschrift, **41** (1936), 345–356.

(4) G. H. Hardy and S. Ramanujan, *Asymptotic formulae in combinatory analysis*, Proc. London Math. Soc. (2), **17** (1918), 75–115 [= Collected works of S. Ramanujan, 276–309 (Paper 36)].

(5) W. T. Martin and N. Wiener, *Taylor's series of functions of smooth growth in the unit circle*, Duke Math. Journal, **4** (1938), 384–392.

(6) H. Rademacher, *On the partition function $p(n)$*, Proc. London Math. Soc. (2), **43** (1937), 241–254.

(7) J. V. Uspensky, *Asymptotic expressions of numerical functions occurring in problems of partition of numbers* [in Russian], Bull. Acad. Sci. de Russie (6), **14** (1920), 199–218.

(8) N. Wiener and H. R. Pitt, *A generalization of Ikehara's theorem*, Journal of Math. and Phys. M.I.T., **17** (1938), 247–258.

(9) E. M. Wright, *Asymptotic partition formulae.* III. *Partitions into k^{th} powers*, Acta Math., **63** (1934), 143–191.

ANNALS OF MATHEMATICS
Vol. 42, No. 5, December, 1941

BEITRÄGE ZUR TOPOLOGIE DER GRUPPEN-MANNIGFALTIGKEITEN

VON HANS SAMELSON

(Received October 18, 1940)

EINLEITUNG

Die Theorie der Lie'schen Gruppen war ursprünglich eine "lokale" Theorie: man betrachtete nur eine Umgebung des Einselementes einer Gruppe, das, was man heute einen Gruppenkeim oder eine lokale Gruppe nennt [1].[1] Diese Einschränkung, die dem Gruppenbegriff eigentlich nicht angemessen ist, wurde in der neueren Forschung aufgegeben. Man betrachtet heute die Algebra der *ganzen* Gruppe und die Geometrie im *Grossen* des die Gruppe repräsentierenden Raumes, der "Gruppen-Mannigfaltigkeit". Die damit auftretenden geometrischen Fragen sind im wesentlichen topologische Probleme. Die hierhergehörigen Ergebnisse, die im Sommer 1935 bekannt waren, hat E. Cartan in einem Vortrag "Sur la topologie des groupes de Lie" zusammengestellt [2]. (In der vorliegenden Arbeit werden nur kompakte Gruppen, deren Räume also geschlossene Mannigfaltigkeiten sind, behandelt.)

Der topologische Bau einer Mannigfaltigkeit wird in erster Linie beschrieben durch ihre Betti'schen Gruppen, sowie durch ihren Homologie-Ring, den man aus den Betti'schen Gruppen bekommt, wenn man die Schnitt-Bildung als Multiplikation einführt.[2] Für *Gruppen*-Mannigfaltigkeiten hat dieser Ring eine besondere Bedeutung: er ist—nach Cartan—isomorph dem Ring der invarianten Differentialformen der Gruppe. Daraus ergibt sich weiter, dass das topologische Problem, die Struktur des Homologie-Ringes einer Gruppe zu bestimmen, äquivalent ist mit einem algebraischen Problem, das von den Invarianten linearer Gruppen handelt. Auf beiden Wegen zur Untersuchung des Ringes sind aber die Schwierigkeiten so gross, dass heute noch nicht die Ringe aller bekannten Gruppen ermittelt sind.

Immerhin ist diese Aufgabe gelöst für die Gruppen, die Weyl die "klassischen" nennt—also die orthogonalen, die unitären und die symplektischen—, und zwar von Pontrjagin auf dem topologischen [3], von R. Brauer auf dem algebraischen Wege [4]. Es stellte sich dabei heraus: Die genannten Gruppen haben die gleichen Homologie-Ringe wie gewisse topologische Produkte von Sphären. Die Dimensionszahlen dieser Sphären, die ja dann den Ring bestimmen, lassen sich für jede dieser Gruppen angeben. Für die Poincaré'schen Polynome—also

[1] Die Nummern im Text verweisen auf das Literaturverzeichnis am Schluss der Arbeit.

[2] Wir verwenden als Koeffizientenbereich in dieser Arbeit immer den Körper der rationalen Zahlen (oder, was fast auf dasselbe herauskommt, den Körper der reellen Zahlen; dieser Körper ist bei der im nächsten Satz genannten Isomorphie mit dem Ring der invarianten Differentialformen benützt).

is monotonic, because $e^{-u\sigma}$ varies only by a constant factor $e^{\pm c}$, when u ranges over an interval of length c/σ; and the sharper conclusion (7) is possible if c can be taken arbitrarily small, because $e^{\pm c} \to 1$ when $c \to 0$.

THE INSTITUTE FOR ADVANCED STUDY
PRINCETON, NEW JERSEY

REFERENCES

(1) V. G. Avakumović, *Über das Verhalten Laplacescher Integrale an der Konvergenzgrenze mit neuem Beweis eines Satzes von Hardy-Ramanujan über das asymptotische Verhalten der Zerfällungskoeffizienten*, (2 Congr. Interbalkan, des Math., Bucarest, 12-IX-1937), Bull. Math. Soc. Roum. Sci. **40**, Nr. 1/2 (1938), 101–106.

(2) V. G. Avakumović, *Bemerkungen über Laplacesche Integrale, deren Wachstum von Exponentialcharakter ist.* I *und* II, Math. Zeitschrift, **46** (1940), 62–69.

(3) V. Avakumović and J. Karamata, *Über einige Taubersche Sätze, deren Asymptotik von Exponentialcharakter ist.* I, Math. Zeitschrift, **41** (1936), 345–356.

(4) G. H. Hardy and S. Ramanujan, *Asymptotic formulae in combinatory analysis*, Proc. London Math. Soc. (2), **17** (1918), 75–115 [= Collected works of S. Ramanujan, 276–309 (Paper 36)].

(5) W. T. Martin and N. Wiener, *Taylor's series of functions of smooth growth in the unit circle*, Duke Math. Journal, **4** (1938), 384–392.

(6) H. Rademacher, *On the partition function* $p(n)$, Proc. London Math. Soc. (2), **43** (1937), 241–254.

(7) J. V. Uspensky, *Asymptotic expressions of numerical functions occurring in problems of partition of numbers* [in Russian], Bull. Acad. Sci. de Russie (6), **14** (1920), 199–218.

(8) N. Wiener and H. R. Pitt, *A generalization of Ikehara's theorem*, Journal of Math. and Phys. M.I.T., **17** (1938), 247–258.

(9) E. M. Wright, *Asymptotic partition formulae.* III. *Partitions into* k^{th} *powers*, Acta Math., **63** (1934), 143–191.

ANNALS OF MATHEMATICS
Vol. 42, No. 5, December, 1941

BEITRÄGE ZUR TOPOLOGIE DER GRUPPEN-MANNIGFALTIGKEITEN

VON HANS SAMELSON

(Received October 18, 1940)

EINLEITUNG

Die Theorie der Lie'schen Gruppen war ursprünglich eine "lokale" Theorie: man betrachtete nur eine Umgebung des Einselementes einer Gruppe, das, was man heute einen Gruppenkeim oder eine lokale Gruppe nennt [1].[1] Diese Einschränkung, die dem Gruppenbegriff eigentlich nicht angemessen ist, wurde in der neueren Forschung aufgegeben. Man betrachtet heute die Algebra der *ganzen* Gruppe und die Geometrie im *Grossen* des die Gruppe repräsentierenden Raumes, der "Gruppen-Mannigfaltigkeit". Die damit auftretenden geometrischen Fragen sind im wesentlichen topologische Probleme. Die hierhergehörigen Ergebnisse, die im Sommer 1935 bekannt waren, hat E. Cartan in einem Vortrag "Sur la topologie des groupes de Lie" zusammengestellt [2]. (In der vorliegenden Arbeit werden nur kompakte Gruppen, deren Räume also geschlossene Mannigfaltigkeiten sind, behandelt.)

Der topologische Bau einer Mannigfaltigkeit wird in erster Linie beschrieben durch ihre Betti'schen Gruppen, sowie durch ihren Homologie-Ring, den man aus den Betti'schen Gruppen bekommt, wenn man die Schnitt-Bildung als Multiplikation einführt.[2] Für *Gruppen*-Mannigfaltigkeiten hat dieser Ring eine besondere Bedeutung: er ist—nach Cartan—isomorph dem Ring der invarianten Differentialformen der Gruppe. Daraus ergibt sich weiter, dass das topologische Problem, die Struktur des Homologie-Ringes einer Gruppe zu bestimmen, äquivalent ist mit einem algebraischen Problem, das von den Invarianten linearer Gruppen handelt. Auf beiden Wegen zur Untersuchung des Ringes sind aber die Schwierigkeiten so gross, dass heute noch nicht die Ringe aller bekannten Gruppen ermittelt sind.

Immerhin ist diese Aufgabe gelöst für die Gruppen, die Weyl die "klassischen" nennt—also die orthogonalen, die unitären und die symplektischen—, und zwar von Pontrjagin auf dem topologischen [3], von R. Brauer auf dem algebraischen Wege [4]. Es stellte sich dabei heraus: Die genannten Gruppen haben die gleichen Homologie-Ringe wie gewisse topologische Produkte von Sphären. Die Dimensionszahlen dieser Sphären, die ja dann den Ring bestimmen, lassen sich für jede dieser Gruppen angeben. Für die Poincaré'schen Polynome—also

[1] Die Nummern im Text verweisen auf das Literaturverzeichnis am Schluss der Arbeit.

[2] Wir verwenden als Koeffizientenberich in dieser Arbeit immer den Körper der rationalen Zahlen (oder, was fast auf dasselbe herauskommt, den Körper der reellen Zahlen; dieser Körper ist bei der im nächsten Satz genannten Isomorphie mit dem Ring der invarianten Differentialformen benützt).

für die Polynome in einer Unbestimmten t, deren Koeffizienten die Betti'schen Zahlen der betreffenden Mannigfaltigkeit sind—folgt, dass sie in Faktoren der Gestalt $(1 + t^m)$ zerfallen; dabei sind die Exponenten m die erwähnten Dimensionszahlen.

Diese wichtigen Untersuchungen von Pontrjagin und Brauer—(ebenso wie eine etwas jüngere topologische Arbeit von Ehresmann [5], in der dieselben Ergebnisse erhalten werden)—haben einen "speziellen" Charakter: es wird jedesmal eine spezielle Klasse von Gruppen betrachtet, und es werden jedesmal ganz spezielle Eigenschaften dieser Gruppen benützt. Daher erhebt sich die Frage nach einer allgemeinen topologischen Theorie der Gruppenmannigfaltigkeiten, aus der man die erwähnten Sätze über die klassischen Gruppen durch Spezialisierung gewinnen könnte. So schliesst auch Cartan seinen genannten Vortrag nach einem Bericht über die Pontrjagin-Brauer'schen Untersuchungen mit den Worten: "Il faut espérer qu'on trouvera aussi une raison de portée générale expliquant la forme si particulière des polynomes de Poincaré des groupes simples clos." [6]

In dieser Hinsicht ist nun ein wichtiger Schritt der folgende Satz von Hopf: [7] "Für jede (geschlossene) Gruppe ist der Homologie-Ring isomorph dem Ring eines gewissen topologischen Produktes von Sphären, deren Dimensionen ungerade sind"; es gilt also allgemein das, was Pontrjagin und Brauer für die klassischen Gruppen festgestellt haben—insbesondere zerfällt das Poincaré'sche Polynom *jeder* Gruppe in Faktoren der Gestalt $(1 + t^m)$ mit ungeraden Exponenten m. Dieser Satz besitzt offenbar die gewünschte Allgemeinheit; jedoch wird in seinem Beweis die Gruppen-Eigenschaft nicht vollständig ausgenützt und vor allem ein wichtiges Hilfsmittel nicht herangezogen, das Pontrjagin in seiner Arbeit [3] eingeführt hat: eine Multiplikation von Zyklen mit Hilfe der Gruppenmultiplikation, die kurz so aussieht: Durchläuft der Punkt p den Zyklus x und der Punkt q den Zyklus y, dann durchläuft der Produktpunkt pq den "Produktzyklus" von x und y.

Auf der Betrachtung dieser "Pontrjagin'schen" Multiplikation und ihrer Kombination mit dem Satz von Hopf basiert nun die vorliegende Arbeit. Und zwar beweisen wir in Kap. I, §3 einen Satz (den "Aufspannsatz") über die Struktur des Ringes einer Gruppe, dessen Beweis für die klassischen Gruppen einen wesentlichen Teil des Inhaltes der Pontrjagin'schen Arbeit bildet und dessen Gültigkeit für alle Gruppen von Cartan [8] vermutet wurde. Der Inhalt dieses Satzes ist eine ganz bestimmte Dualität zwischen der Gruppen- (Pontrjagin'schen)-Multiplikation der Zyklen und der Schnitt-Bildung der Zyklen; dadurch wird die Gruppe in eine noch engere Analogie zu einem Sphären-Produkt gebracht als durch den Satz von Hopf; es handelt sich um einen "doppelten" Isomorphismus.

Der §1 von Kap. I enthält eine Zusammenstellung bekannter Tatsachen über Homologie-Ringe von Mannigfaltigkeiten, und die Definition und Eigenschaften der (bei Hopf [7] eingeführten) Begriffe "maximales" und "minimales" Element eines Homologie-Ringes. In §2 wird der Satz von Hopf zitiert, und

es werden die für das folgende grundlegenden Eigenschaften des Ringes einer Gruppe dargestellt. Der §3 enthält den Beweis des oben genannten "Aufspannsatzes".

Auf Grund dieses Satzes werden dann in Kap. II die beiden wichtigsten gruppentheoretischen Begriffe "homomorphe Abbildung" und "Untergruppe" in topologischer Hinsicht untersucht. Es wird die Frage diskutiert, wie die Mannigfaltigkeit einer Untergruppe in die Gruppe eingelagert ist, d.h. welche Zyklen aus der Untergruppe in der Gruppe homolog Null sind; dabei ergeben sich einfache Gesetze. Diese spielen im folgenden eine Rolle.

Eine kontinuierliche Gruppe G ist gewöhnlich gegeben als transitive Transformationsgruppe einer Mannigfaltigkeit W, des "Wirkungsraumes" ("espace homogène") [9]. Diejenigen Transformationen, die einen bestimmten Punkt von W festhalten, bilden eine Untergruppe U von G, die sogenannte "Isotropiegruppe" von W, die bis auf innere Automorphismen von G unabhängig von der Wahl des festen Punktes ist. Die Nebengruppen, in welche G nach U zerfällt, stehen in eineindeutiger und stetiger Beziehung zu den Punkten von W; geometrisch gesehen liegt eine "Faserung" von G in die (untereinander homöomorphen) Nebengruppen von U vor [10].

Eines der Hauptprobleme der weiteren Theorie besteht nun in der Untersuchung der Beziehungen zwischen G, U und W. Einerseits ist das der natürliche Ansatzpunkt zur Untersuchung einer als Transformationsgruppe gegebenen Gruppe; andererseits sind die Wirkungsräume, also die Mannigfaltigkeiten, die transitive Transformationsgruppen zulassen, an und für sich interessant. Hier hat man nun Möglichkeiten, wenn man die Sätze der Kap. I und II dieser Arbeit verknüpft mit topologischen Sätzen über die Faserungen von Mannigfaltigkeiten. Dazu werden in Kap. III und IV einige Beiträge geliefert.

Das Kap. III behandelt den Fall, dass der Wirkungsraum W eine Sphäre ist. Das Ergebnis ist, dass man auf eine sehr einfache Weise aus dem Homologie-Ring der Isotropiegruppe U und der Dimension der Wirkungs-Sphäre W den Ring der Gruppe G bestimmen kann. Dieser allgemeine Satz gibt mühelos durch eine Rekursion die Pontrjagin-Brauer'schen Resultate über die Ringe der "klassischen" Gruppen; denn diese Gruppen sind bekanntlich transitive Transformationsgruppen von Sphären, und die zugehörigen Isotropiegruppen sind auch klassische Gruppen.

Im Kap. IV werden die Beziehungen zwischen G, U und W ohne spezielle Voraussetzungen über W untersucht. Unter anderem wird der Satz bewiesen: Ist U nicht homolog 0 in G, dann ist der Homologie-Ring von G isomorph dem Ring des topologischen Produktes von U und W. Dabei ergibt sich noch: Ein Wirkungsraum W, dessen Isotropiegruppe U nicht homolog 0 in der Gruppe G ist, hat den gleichen Ring wie ein gewisses Produkt von Sphären ungerader Dimensionen. Das ist eine Verallgemeinerung des Satzes von Hopf über den Ring einer Gruppen-Mannigfaltigkeit (eine Gruppe lässt sich auffassen als Wirkungsraum, dessen Isotropiegruppe ein Punkt ist, und ein Punkt ist nie homolog 0).

Kapitel I

Ein Satz über Gruppen-Mannigfaltigkeiten

Die §§1 und 2 enthalten eine geeignete Zusammenstellung bekannter Dinge (im wesentlichen ohne Beweise) und einen Bericht über eine Arbeit von Hopf über Gruppenmannigfaltigkeiten, auf die sich die vorliegende Arbeit stützt. In §3 wird ein allgemeiner Satz über Gruppenmannigfaltigkeiten bewiesen.

1. Der Homologie-Ring einer Mannigfaltigkeit. Irreduzible Erzeugendensysteme, maximale und minimale Elemente

1. M sei eine geschlossene, orientierbare Mannigfaltigkeit der Dimension n. Die Homologieklassen von M, die Elemente der Betti'schen Gruppe $\mathfrak{B}(M)$, bilden einen Ring, den Homologie- oder Schnittring $\mathfrak{R}(M)$, wenn man die Addition durch die Addition in der Betti'schen Gruppe und das Produkt durch die Schnittbildung erklärt [11]. *Dabei sollen in dieser Arbeit immer die rationalen Zahlen als Koeffizienten dienen.*

$\mathfrak{R}(M)$ enthält ein Eins-Element, nämlich den orientierten n-dimensionalen Grundzyklus von M; es wird mit 1 bezeichnet. Die Dimension eines homogen-dimensionalen Elementes z von $\mathfrak{R}(M)$ bezeichenen wir mit $d(z)$; die Zahl $\delta(z) = n - d(z)$ heisse die duale Dimension von z. Das Null-Element 0 hat jede Dimension.

Für homogen-dimensionale Elemente z, z' ist auch das Produkt oder der Schnitt $z \cdot z'$ homogen-dimensional und zwar gilt:

$$\delta(z \cdot z') = \delta(z) + \delta(z').$$

Weiter ist:

$$z \cdot z' = (-1)^{\delta(z)\cdot\delta(z')} z' \cdot z,$$

also speziell:

$$z \cdot z = 0 \qquad \text{bei ungeradem } \delta(z).$$

Die rationalen Vielfachen der 1 heissen die Skalare.

Wo im folgenden Homologieklassen auftreten, sollen sie immer, wenn nicht anders bemerkt, homogen-dimensional sein.

2. $\mathfrak{B}_r(M)$, oder kurz $\mathfrak{B}_r$, sei die r-te Betti'sche Gruppe von M, also die additive Gruppe der r-dimensionalen Elemente von $\mathfrak{R}(M)$; ihr Rang p_r ist die r-te Betti'sche Zahl von M. Die volle Betti'sche Gruppe $\mathfrak{B}(M)$ ist die direkte Summe

$$\mathfrak{B}_0(M) + \mathfrak{B}_1(M) + \cdots + \mathfrak{B}_n(M).$$

Für $r < n$ sei $\mathfrak{U}_r$ die Gruppe derjenigen Elemente von $\mathfrak{B}_r$, die sich aus den Elementen der Gruppen

$$\mathfrak{B}_{n-1}, \mathfrak{B}_{n-2}, \cdots, \mathfrak{B}_{r+1}$$

durch Multiplikation und Addition erzeugen lassen (vgl. [7], Nr. 30), die sich also als $\sum x_h \cdot y_h$ mit $\delta(x_h) > 0$ und $\delta(y_h) > 0$ schreiben lassen; die Elemente von $\mathfrak{U}_r$ heissen "zusammengesetzt."

Ein nicht-skalares Element heisst maximal, wenn es nicht zusammengesetzt ist, wenn es sich also nicht als Summe von Produkten von höherdimensionalen Elementen darstellen lässt. Ist z maximal und w zusammengesetzt, dann ist auch $z + w$ maximal, weil $\mathfrak{U}_r$ eine Gruppe ist. Die Skalare werden nicht als maximal angesehen; die höchstdimensionalen unter den nicht-skalaren, nicht-verschwindenden Elementen sind immer maximal.

Man stelle $\mathfrak{B}_r$ als direkte Summe $\mathfrak{U}_r + \mathfrak{Z}_r$ dar; das ist möglich, weil die rationalen Zahlen als Koeffizienten dienen. Die Elemente von $\mathfrak{Z}_r$ sind dann maximal. Wählt man in jeder Gruppe $\mathfrak{Z}_r$ eine Basis, so bildet die Gesamtheit dieser Basiselemente mit der 1 zusammen, wie man leicht sieht, ein Erzeugendensystem des Ringes $\mathfrak{R}(M)$; d.h.: jedes Element von $\mathfrak{R}(M)$ lässt sich als Polynom in den Elementen dieses Systems schreiben. Man bestätigt auch leicht, dass ein irreduzibles Erzeugendensystem vorliegt, d.h. dass kein echtes Teilsystem ein Erzeugendensystem ist.

Der Rang der Gruppe $\mathfrak{Z}_{n-r}$, also auch der Rang der Restklassengruppe $\mathfrak{B}_{n-r} - \mathfrak{U}_{n-r}$, werde l_r genannt. (Rang = Maximalzahl linear unabhängiger Elemente.) Man stellt fest, dass die nicht-skalaren Elemente eines beliebigen irreduziblen Erzeugendensystems $(1, z_1, z_2, \cdots, z_l)$ von $\mathfrak{R}(M)$ maximal sind, und dass die Zahl derjenigen z_i, für die $\delta(z_i) = r$ ist, immer gleich l_r ist; die Anzahl der Elemente eines solchen Systems (ohne die 1) ist also immer

$$l = l_1 + l_2 + \cdots + l_n . \qquad \text{([7], Nr. 31)}$$

Ersetzt man die Elemente z_i eines solchen Systems durch $z_i + y_i$, wo die Elemente y_i nicht maximal sind, so erhält man wieder ein irreduzibles Erzeugendensystem; das kommt daher, dass die z_i Repräsentanten für die Restklassengruppe $\mathfrak{B}_r - \mathfrak{U}_r$ sind.

3. Unter den Vielfachen $x \cdot v$ eines Elementes v von $\mathfrak{R}(M)$, wo x alle Elemente von $\mathfrak{R}(M)$ durchläuft, treten immer auf:

a) die rationalen Vielfachen von v, deren Dimension gleich der von v ist—man setze für x einen beliebigen Skalar ein—und

b) alle null-dimensionalen Elemente, d.h. der Punkt mit einem beliebigen Koeffizienten, wenn $v \neq 0$ ist—wegen des Poincaré-Veblen'schen Dualitätssatzes.

Wir definieren nun:

Ein Element v von $\mathfrak{R}(M)$ heisst minimal, wenn es keine anderen Vielfachen besitzt, d. h. wenn aus $0 < d(x \cdot v) < d(v)$ folgt: $x \cdot v = 0$. (Statt $0 < d(x \cdot v) < d(v)$ kann man auch sagen: $0 < \delta(x) < d(v)$.) (Vgl. [7], Nr. 32)

Die nulldimensionalen Elemente gelten nicht als minimal, so wie die n-dimensionalen Elemente nicht als maximal gelten. Dagegen ist die Null minimal.

Es gilt der folgende Satz:

Das s-dimensionale Element v ist dann und nur dann minimal, wenn es ein

Annulator der Gruppe $\mathfrak{U}_{n-s}$ *ist, d.h. wenn* $u \cdot v = 0$ *ist für jedes Element* u *von* $\mathfrak{U}_{n-s}$ $(s = 1, \cdots, n)$ ([7], Nr. 33). Daraus entnimmt man: die s-dimensionalen minimalen Elemente bilden eine additive Gruppe $\mathfrak{B}_s$, eine Untergruppe von $\mathfrak{B}_s$. Aus dem Poincaré-Veblen'schen Dualitätssatz folgt leicht, dass der Rang von $\mathfrak{B}_s$ gleich l_s ist, nämlich gleich dem Rang von $\mathfrak{Z}_{n-s}$ in der direkten Summe $\mathfrak{B}_{n-s} = \mathfrak{U}_{n-s} + \mathfrak{Z}_{n-s}$; $\mathfrak{B}_s$ und $\mathfrak{Z}_{n-s}$ sind gewissermassen dual zueinander. Die direkte Summe $\mathfrak{B}_1 + \mathfrak{B}_2 + \cdots + \mathfrak{B}_n$ aller dieser Gruppen heisst die Gruppe der minimalen Elemente von $\mathfrak{R}(M)$; sie wird mit $\mathfrak{B}(M)$, oder kurz mit $\mathfrak{B}$ bezeichnet. Ihr Rang ist

$$l = l_1 + l_2 + \cdots + l_n .$$

4. M und M' seien zwei Mannigfaltigkeiten, f eine Abbildung von M in M'.[3] f bildet den Ring $\mathfrak{R}(M)$ in den Ring $\mathfrak{R}(M')$ ab. Dann gilt der wichtige Invarianzsatz:

Ist v *ein minimales Element von* $\mathfrak{R}(M)$, *so ist sein Bild* $f(v)$ *ein minimales Element von* $\mathfrak{R}(M')$. ([7], Nr. 34)

Eine spezielle Folgerung hieraus ist:

Ein Element von $\mathfrak{R}(M)$, das durch das stetige Bild einer Sphäre S_n $(n > 0)$ repräsentiert wird, ist ein minimales Element. Denn in $\mathfrak{R}(S_n)$ ist der Grundzyklus S_n—die 1 von $\mathfrak{R}(S_n)$—minimal.

5. M und M' seien zwei geschlossene, orientierbare Mannigfaltigkeiten; f sei eine Abbildung von M in M'. Durch f wird eine Abbildung von $\mathfrak{R}(M)$ in $\mathfrak{R}(M')$ bewirkt, die auch f heisse; sie ist additiv, aber im allgemeinen nicht multiplikativ homomorph. Dann existiert eine Abbildung φ von $\mathfrak{R}(M')$ in $\mathfrak{R}(M)$—der Umkehrungshomomorphismus—mit folgenden Eigenschaften: [12]

1) φ ist ein Ring-Homomorphismus, also additiv und multiplikativ homomorph,

2) φ ist mit f durch die Funktionalgleichung

$$f(\varphi(z') \cdot x) = z' \cdot f(x)$$

verknüpft, in der x ein beliebiges Element von $\mathfrak{R}(M)$ und z' ein beliebiges Element von $\mathfrak{R}(M')$ ist,

3) ist z' ein homogen-dimensionales Element von $\mathfrak{R}(M')$, so ist auch $\varphi(z')$ homogen-dimensional, und zwar ist

$$\delta(\varphi(z')) = \delta(z'),$$

anders geschrieben:

$$d(\varphi(z')) = d(z') + d(M) - d(M').$$

(Vgl. [7], Nr. 25) Mit Hilfe dieser Eigenschaften lässt sich allgemein bei gegebener Abbildung f, bzw. φ, die Abbildung φ, bzw. f, anschreiben; wir werden

[3] Alle in dieser Arbeit vorkommenden Abbildungen sind eindeutig und stetig.

jedoch diese Beziehung, die wir nur in spezieller Weise brauchen, jedesmal direkt ableiten.

6. Es soll jetzt noch ein Hilfssatz abgeleitet werden, der später gebraucht wird. Es sei z_0, bzw. z_0', der einfach gezählte Punkt von M, bzw. M'.

HILFSSATZ: *Ist* $\varphi(z_0') \neq 0$, *so ist* f *eine Abbildung von* $\mathfrak{R}(M)$ ***auf*** $\mathfrak{R}(M')$.

Es sei also $\varphi(z_0') \neq 0$; wir zeigen zunächst: Ist x' von Null verschieden, dann ist auch $\varphi(x')$ nicht Null. Weil x' nicht Null ist, gibt es ein duales y' mit $y' \cdot x' = z_0'$. Wegen der Multiplikativität von φ ist $\varphi(y') \cdot \varphi(x') = \varphi(y' \cdot x') = \varphi(z_0') \neq 0$. Also ist auch $\varphi(x') \neq 0$. Die Gruppe derjenigen Elemente von $\mathfrak{B}(M')$, die bei der Abbildung f als Bilder von Elementen von $\mathfrak{B}(M)$ auftreten, die Bildgruppe $f(\mathfrak{B}(M))$, hat dann die folgende Eigenschaft:

Sie enthält zu einem beliebigen Element $x'(\neq 0)$ von $\mathfrak{B}(M')$ ein Element y' mit $x' \cdot y' = z_0'$.

Denn aus $x' \neq 0$ folgt $\varphi(x') \neq 0$; also gibt es ein y so, dass $\varphi(x') \cdot y = z_0$ ist; übt man auf diese Gleichung f aus und beachtet die Funktionalgleichung für φ, so erhält man $x' \cdot f(y) = z_0'$. Die Behauptung ist also mit $y' = f(y)$ erfüllt.

Aber aus dieser Eigenschaft der Bildgruppe folgt nach dem Poincaré-Veblen'schen Dualitätssatz sofort, dass die Bildgruppe mit $\mathfrak{B}(M')$ zusammenfällt, was zu beweisen war.

7. Zu zwei Mannigfaltigkeiten M und M' gehört die Produktmannigfaltigkeit $M \times M'$. Ueber die Homologieeigenschaften von $M \times M'$ ist bekannt [13] (vgl. [7], Nr. 19):

Zu jedem Element x von $\mathfrak{R}(M)$ und jedem Element x' von $\mathfrak{R}(M')$ gehört ein Element $x \times x'$ von $\mathfrak{R}(M \times M')$; diese Produktbildung ist mit der Addition distributiv verknüpft; es gilt: $d(x \times x') = d(x) + d(x')$.

Jedes Element von $\mathfrak{R}(M \times M')$ lässt sich als $\sum (x_h \times x_h')$ darstellen; genauer gilt: Ist $(Z_1, Z_2, \cdots Z_p)$ eine additive Basis von $\mathfrak{R}(M)$—d.h. eine Basis von $\mathfrak{B}(M)$—und ist $(Z_1', Z_2', \cdots Z_{p'}')$ eine additive Basis von $\mathfrak{R}(M')$, so bilden die Elemente $Z_i \times Z_k'$ eine additive Basis in $\mathfrak{R}(M \times M')$; die Elemente von $\mathfrak{R}(M \times M')$ lassen sich in eineindeutiger Weise als $\sum t_{ik}(Z_i \times Z_k')$ mit rationalen Koeffizienten t_{ik} schreiben. (All das gilt auch für den Produktkomplex zweier Komplexe.)

Für die Multiplikation gilt [14]:

$$(x \times x') \cdot (y \times y') = (-1)^{\delta(x) \cdot \delta(y')} (x \cdot y \times x' \cdot y');$$

dabei müssen x und y' homogen-dimensional sein. Ausserdem gelten die üblichen distributiven Gesetze.

8. Der folgende Abbildungssatz für topologische Produkte wird später öfters verwendet:

A, B, A', B' seien Mannigfaltigkeiten; f sei eine Abbildung von A in A', g sei eine Abbildung von B in B'. Man konstruiert eine Abbildung h von $A \times B$ in $A' \times B'$, indem man setzt:

$$h(p \times q) = f(p) \times g(q).$$

Ist dann x, bzw. y, ein Homologieelement von A, bzw. B, so gilt:

$$h(x \times y) = f(x) \times g(y).$$

Man beweist das, indem man f und g simplizial approximiert. Ist dann a, bzw. b, ein Simplex von A, bzw. B, so ist das Bild der Zelle $a \times b$ die Zelle $f(a) \times g(b)$:

$$h(a \times b) = f(a) \times g(b);$$

denn beide Zellen rechts und links haben dieselben Eckpunkte, und auch die Orientierung stimmt überein, wie man leicht sieht. Durch Addition gewinnt man die Formel dann für Komplexe, Zyklen und Homologieklassen.

9. Wie man ganz leicht sieht, besteht die Gruppe $\mathfrak{B}(M \times M')$ der minimalen Elemente von $M \times M'$ genau aus den Elementen $(v \times z_0') + (z_0 \times v')$, wo z_0, bzw. z_0', der einfach gezählte Punkt von M, bzw. M', ist und v, bzw. v', ein beliebiges Element von $\mathfrak{B}(M)$, bzw. $\mathfrak{B}(M')$, ist.

2. Der Ring einer Gruppen-Mannigfaltigkeit

1. G sei eine geschlossene Gruppen-Mannigfaltigkeit, also eine geschlossene Mannigfaltigkeit (der Dimension n), zwischen deren Punkten eine Multiplikation erklärt ist, die die bekannten Gruppenaxiome erfüllt, und die die weitere Eigenschaft hat, dass das Produkt pq zweier Punkte p, q von G stetig von dem Paar (p, q) abhängt, und ebenso das Inverse p^{-1} stetig von p abhängt. (Gruppenmannigfaltigkeiten sind bekanntlich orientierbar.)

Ueber den Homologie-Ring von G ist der folgende Satz bekannt ([7], Nr. 2):

Satz A. *Der Ring $\mathfrak{R}(G)$ (mit rationalen Koeffizienten) einer Gruppen-Mannigfaltigkeit G ist dimensionstreu isomorph dem Ring $\mathfrak{R}(\Pi)$ eines topologischen Produktes*

$$\Pi = S_{m_1} \times S_{m_2} \times \cdots \times S_{m_l}, \qquad l \geqq 1,$$

wo S_m die m-dimensionale Sphäre bezeichnet, und die Dimensionen m_i ungerade sind.

2. Der Beweis dieses Satzes verläuft so, dass man zeigt: Bilden die Elemente $z_1, z_2, \cdots z_l$ mit der 1 zusammen ein irreduzibles Erzeugendensystem von $\mathfrak{R}(G)$, dann ist

a) das Produkt $z_1 \cdot z_2 \cdot \cdots z_l$ von Null verschieden,

b) $\delta(z_i) = m_i$ ungerade $(i = 1, \cdots, l)$ ([7], Nr. 15).

Daraus folgert man: Es ist

$$z_i \cdot z_k = -z_k \cdot z_i, \quad \text{speziell} \quad z_i \cdot z_i = 0;$$

und eine volle additive Basis von $\mathfrak{R}(G)$ (eine Basis von $\mathfrak{B}(G)$) wird gebildet von den 2^l Elementen

$$1, z_i, z_{i_1}\cdot z_{i_2}\ (i_1 < i_2), z_{i_1}\cdot z_{i_2}\cdot z_{i_3}\ (i_1 < i_2 < i_3), \cdots, z_1\cdot z_2\cdot \cdots \cdot z_l .$$

Diese Basis, die also durch das irreduzible Erzeugendensystem $(1, z_1, z_2, \cdots, z_l)$ bestimmt ist, heisse die Basis $[z_i]$.

Sind x, y zwei Elemente ($\neq 1$) dieser Basis, so ist $x\cdot y \neq 0$, wenn kein Element z_i in beiden Elementen zugleich als Faktor auftritt, dagegen ist $x\cdot y = 0$, wenn wenigstens ein z_i in beiden Elementen als Faktor auftritt, wie man leicht mit Hilfe der Assoziativität, der Antikommutativität der Elemente z_i und der Tatsache $z_1\cdot z_2\cdot \cdots \cdot z_l \neq 0$ bestätigt. Man kann $\mathfrak{R}(G)$ als den Ring der inhomogenen Multilinearformen in den antikommutativen Grössen $z_1, z_2, \cdots z_l$ beschreiben. Die Dimension von G ist

$$n = m_1 + m_2 + \cdots + m_l .$$

Das Produkt $z_1\cdot z_2\cdot \cdots \cdot z_l$ ist nulldimensional; man kann annehmen, dass es gleich dem einfach gezählten Punkt von G ist, der v_0 heissen soll (man kann nämlich etwa z_1 durch ein geeignetes rationales Vielfaches ersetzen).

Die Isomorphie zwischen $\mathfrak{R}(G)$ und $\mathfrak{R}(\Pi)$ wird folgendermassen hergestellt: Der orientierte Grundzyklus der Sphäre S_m werde auch mit S_m bezeichnet; p sei ein einfach gezählter Punkt von S_m. Man bilde die Elemente

$$\begin{aligned} Z_1 &= p \times S_{m_2} \times S_{m_3} \times \cdots \times S_{m_l} \\ Z_2 &= S_{m_1} \times p \times S_{m_3} \times \cdots \times S_{m_l} \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ Z_l &= S_{m_1} \times S_{m_2} \times S_{m_3} \times \cdots \times p \end{aligned}$$

von $\mathfrak{R}(\Pi)$. Die Z_i haben auch die Eigenschaften a) und b); die Zuordnung $z_i \leftrightarrow Z_i$ ergibt offenbar einen Isomorphismus von $\mathfrak{R}(G)$ und $\mathfrak{R}(\Pi)$.

Die Elemente z_i sind maximal, weil sie zu einem irreduziblen Erzeugendensystem gehören (§1, Nr. 2).

Ein Produkt $z_{i_1}\cdot z_{i_2}\cdot \cdots \cdot z_{i_r}$ ist durch seine Faktoren bis auf das Vorzeichen bestimmt; in diesem Sinne kann man von *dem* Produkt der z_{i_k} sprechen, ohne auf ihre Reihenfolge zu achten.

3. Die Gruppe $\mathfrak{B}(G)$ der minimalen Elemente von $\mathfrak{R}(G)$ lässt sich leicht bestimmen. Von den Elementen der Basis $[z_i]$ sind minimal die und nur die, die Produkte von $l-1$ Faktoren z_i sind, also die Elemente

$$v_i = z_1\cdot z_2\cdot \cdots \cdot z_{i-1}\cdot z_{i+1}\cdot \cdots \cdot z_l .$$

Denn ist x ein Element dieser Basis, das Produkt von höchstens $l-2$ Faktoren z_i ist, also mindestens zwei der z_i, etwa z_r und z_s, nicht als Faktoren enthält, so ist $z_r\cdot x \neq 0$ und $0 < d(z_r\cdot x) < d(x)$, also x nicht minimal. Dagegen ist $y\cdot v_i = 0$ für jedes von z_i und 1 verschiedene Basiselement y, weil y dann mindestens einen Faktor z_k enthält, der auch in v_i vorkommt; aus $0 < \delta(y) < d(v_i)$ folgt also $y\cdot v_i = 0$ (denn wegen $\delta(z_i) = d(v_i)$ ist dann $y \neq z_i$); also ist v_i minimal. Mit derselben Ueberlegung stellt man fest, dass die von den Elementen v_i

erzeugte Untergruppe von $\mathfrak{B}(G)$ schon die ganze Gruppe $\mathfrak{B}(G)$ ist, dass also die minimalen Elemente von $\mathfrak{R}(G)$ genau die Linearkombinationen $\sum a_i v_i$ mit rationalen Koeffizienten a_i sind; die Elemente v_i $(i = 1, \cdots, l)$ bilden also eine Basis von $\mathfrak{B}(G)$.

Der Schnitt $z_i \cdot v_i$ ist gleich $\pm v_0$. Wir wollen das Vorzeichen der Elemente v_i so definieren, dass $z_i \cdot v_i = v_0$ gilt; die v_i erfüllen dann die Gleichungen

$$z_i \cdot v_k = \delta_{ik} v_0 \qquad (i, k = 1, \cdots, l),$$

durch die sie, wie man leicht überlegt, auch vollständig bestimmt sind.

Im Sphärenprodukt Π entsprechen den v_i die "Sphären selbst," nämlich die Elemente

$$\begin{aligned} V_1 &= S_{m_1} \times p \times p \times \cdots \times p \\ V_2 &= p \times S_{m_2} \times p \times \cdots \times p \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ V_l &= p \times p \times p \times \cdots \times S_{m_l}, \end{aligned}$$

die offenbar eine Basis von $\mathfrak{B}(\Pi)$ darstellen. (Um die richtigen Vorzeichen zu erhalten, muss man die Elemente Z_i aus Nr. 2 durch die Elemente $(-1)^{l-i} Z_i$ ersetzen; es gelten nämlich die Gleichungen $Z_i \cdot V_k = (-1)^{l-i} \delta_{ik} V_0$ mit $V_0 = p \times p \times \cdots \times p$. Die in Nr. 2 genannten Eigenschaften der Z_i gelten auch für die Elemente $(-1)^{l-i} Z_i$.)

Hieran sieht man unmittelbar, dass $\sum d(v_i) = n$ ist (das gilt dann natürlich für jede Basis $(v_1', v_2', \cdots v_l')$ von $\mathfrak{B}(G)$).

Die Dimension $d(v_i) = \delta(z_i) = d(S_{m_i}) = m_i$ ist ungerade; in einer Gruppe gibt es also nur ungerade-dimensionale minimale Elemente ([7], Nr. 36).

Die Zahlen l_i, die Ränge der Gruppen $\mathfrak{B}_i(G)$ (§1, Nr. 3), geben an, wieviel der Faktor-Sphären S_{m_k} des Produktes Π i-dimensional sind; durch die l_i ist Π, und folglich auch $\mathfrak{R}(G)$ vollständig bestimmt. Die Zahlen l_i sollen die charakteristischen Zahlen der Gruppe G genannt werden. l_i ist Null für gerades i. Die Zahl

$$l(G) = l_1 + l_3 + l_5 + \cdots$$

soll der Rang[4] der Gruppe heissen; der Rang einer Gruppe G ist also gleich dem Rang der Gruppe $\mathfrak{B}(G)$ der minimalen Elemente von G.

4. Die von den Elementen

$$z_{i_1} \cdot z_{i_2} \cdot \cdots \cdot z_{i_r} \quad (i_1 < i_2 < \cdots < i_r), \quad 1$$

gebildete Basis $[z_i]$ ist zu sich selbst dual, d.h. die duale Basis besteht aus denselben Elementen (bis auf das Vorzeichen).

---

[4] Für den Zusammenhang dieser Definition des Ranges einer Gruppe mit der üblichen Definition des Ranges einer Lieschen Gruppe vgl.: H. Hopf: *Über den Rang geschlossener Liescher Gruppen* [Comm. Math. Helv. 13 (1940/41)].

Ist nämlich $x = z_{i_1} \cdot z_{i_2} \cdot \cdots \cdot z_{i_r}$ ein beliebiges dieser Elemente, und ist $\bar{x} = z_{j_1} \cdot z_{j_2} \cdot \cdots \cdot z_{j_{l-r}}$ das Produkt derjenigen z_i, die in x nicht als Faktoren auftreten, so ist $\bar{x} \cdot x = \pm v_0$; ist aber y ein von $\bar{x}$ verschiedenes Element der Basis $[z_i]$ mit $d(y) = d(\bar{x})$, so ist $y \cdot x = 0$, weil y mindestens einen Faktor z_i enthält, der auch in x vorkommt.

In Nr. 2 sind wir ausgegangen von einem beliebigen irreduziblen Erzeugendensystem $(z_1, z_2, \cdots, z_l, 1)$ von $\mathfrak{R}(G)$; nach Nr. 3 sind den z_i die minimalen Elemente v_i zugeordnet, die eine Basis von $\mathfrak{B}(G)$ bilden. Man sieht ganz leicht, dass man auf diese Weise, d.h. durch geeignete Wahl der z_i, jede Basis von $\mathfrak{B}(G)$ erhalten kann (mit der unwesentlichen Einschränkung, dass man, wegen der Bedingung $z_1 \cdot z_2 \cdot \cdots \cdot z_l = v_0$, nur solche Basen von $\mathfrak{B}(G)$ bekommt, die durch eine lineare Transformation mit der Determinante 1 auseinander hervorgehen; das bedeutet nur, dass man eventuell v_1 durch ein rationales Vielfaches ersetzen muss.). Das beweist man z.B. mit Hilfe der Bemerkung, dass die Gleichungen $z_i \cdot v_k = \delta_{ik} v_0$, durch die bei gegebenen z_i die v_i vollständig bestimmt sind, invariant bleiben, wenn man auf die z_i und die v_i kontragrediente lineare Transformationen ausübt.

Man nehme die zu einem beliebigen irreduziblen Erzeugendensystem $(z_1, z_2, \cdots, z_l, 1)$ von $\mathfrak{R}(G)$ gehörenden minimalen Elemente v_i in eine sonst beliebige Basis von $\mathfrak{B}(G)$ auf. Dann ist das zu v_i duale Element z_i' der dualen Basis von der Form $z_i' = z_i + y_i$ mit nichtmaximalem y_i; das bestätigt man leicht auf Grund der Tatsache, dass $z_i \cdot v_i = v_0$ ist, dass aber für jedes von z_i und 1 verschiedene Element x der Basis $[z_i]$ der Schnitt $x \cdot v_i$ gleich Null ist. Diese Elemente z_i' bilden ebenfalls mit 1 zusammen ein irreduzibles Erzeugendensystem von $\mathfrak{R}(G)$; sie besitzen auch die Eigenschaften a) und b) (§1, Nr. 2); es ist nämlich $\delta(z_i') = \delta(z_i)$ und $z_1' \cdot z_2' \cdot \cdots \cdot z_l' = z_1 \cdot z_2 \cdot \cdots \cdot z_l$; aus den Gleichungen $z_i \cdot v_k = \delta_{ik} v_0$ folgen ohne weiteres die Gleichungen $z_i' \cdot v_k = \delta_{ik} v_0$.

3. Der Pontrjagin'sche Ring und der "Aufspann-Satz"

1. G sei, wie in §1, eine (geschlossene) Gruppenmannigfaltigkeit. Die Multiplikation in G, die jedem Punktepaar (p, q) von G den Punkt $p \cdot q$ zuordnet, kann man auffassen als eine Abbildung F des topologischen Produktes $G \times G$ zweier Exemplare von G in G, die definiert ist durch:

$$F(p \times q) = p \cdot q.$$

Sind x, y zwei Homologieklassen von G, so wollen wir unter dem Pontrjagin'schen Produkt [3] *(kürzer: P-Produkt) von x und y das Bild des Elementes $x \times y$ von $\mathfrak{R}(G \times G)$ bei F verstehen; wir bezeichnen es mit $x \circ y$; also:*

$$F(x \times y) = x \circ y.$$

Anschaulich, aber etwas unscharf, kann man das so sagen: Durchläuft der Punkt p den Zyklus x, und der Punkt q den Zyklus y, dann durchläuft der Punkt pq den Zyklus $x \circ y$.

Es ist $d(x \circ y) = d(x) + d(y)$.

erzeugte Untergruppe von $\mathfrak{B}(G)$ schon die ganze Gruppe $\mathfrak{B}(G)$ ist, dass also die minimalen Elemente von $\mathfrak{R}(G)$ genau die Linearkombinationen $\sum a_i v_i$ mit rationalen Koeffizienten a_i sind; die Elemente v_i $(i = 1, \cdots, l)$ bilden also eine Basis von $\mathfrak{B}(G)$.

Der Schnitt $z_i \cdot v_i$ ist gleich $\pm v_0$. Wir wollen das Vorzeichen der Elemente v_i so definieren, dass $z_i \cdot v_i = v_0$ gilt; die v_i erfüllen dann die Gleichungen

$$z_i \cdot v_k = \delta_{ik} v_0 \qquad (i, k = 1, \cdots, l),$$

durch die sie, wie man leicht überlegt, auch vollständig bestimmt sind.

Im Sphärenprodukt Π entsprechen den v_i die "Sphären selbst," nämlich die Elemente

$$\begin{aligned} V_1 &= S_{m_1} \times p \times p \times \cdots \times p \\ V_2 &= p \times S_{m_2} \times p \times \cdots \times p \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ V_l &= p \times p \times p \times \cdots \times S_{m_l}, \end{aligned}$$

die offenbar eine Basis von $\mathfrak{B}(\Pi)$ darstellen. (Um die richtigen Vorzeichen zu erhalten, muss man die Elemente Z_i aus Nr. 2 durch die Elemente $(-1)^{l-i} Z_i$ ersetzen; es gelten nämlich die Gleichungen $Z_i \cdot V_k = (-1)^{l-i} \delta_{ik} V_0$ mit $V_0 = p \times p \times \cdots \times p$. Die in Nr. 2 genannten Eigenschaften der Z_i gelten auch für die Elemente $(-1)^{l-i} Z_i$.)

Hieran sieht man unmittelbar, dass $\sum d(v_i) = n$ ist (das gilt dann natürlich für jede Basis $(v_1', v_2', \cdots v_l')$ von $\mathfrak{B}(G)$).

Die Dimension $d(v_i) = \delta(z_i) = d(S_{m_i}) = m_i$ ist ungerade; in einer Gruppe gibt es also nur ungerade-dimensionale minimale Elemente ([7], Nr. 36).

Die Zahlen l_i, die Ränge der Gruppen $\mathfrak{B}_i(G)$ (§1, Nr. 3), geben an, wieviel der Faktor-Sphären S_{m_k} des Produktes Π i-dimensional sind; durch die l_i ist Π, und folglich auch $\mathfrak{R}(G)$ vollständig bestimmt. Die Zahlen l_i sollen die charakteristischen Zahlen der Gruppe G genannt werden. l_i ist Null für gerades i. Die Zahl

$$l(G) = l_1 + l_3 + l_5 + \cdots$$

soll der Rang[4] der Gruppe heissen; der Rang einer Gruppe G ist also gleich dem Rang der Gruppe $\mathfrak{B}(G)$ der minimalen Elemente von G.

4. Die von den Elementen

$$z_{i_1} \cdot z_{i_2} \cdot \cdots \cdot z_{i_r} \quad (i_1 < i_2 < \cdots < i_r), \quad 1$$

gebildete Basis $[z_i]$ ist zu sich selbst dual, d.h. die duale Basis besteht aus denselben Elementen (bis auf das Vorzeichen).

---

[4] Für den Zusammenhang dieser Definition des Ranges einer Gruppe mit der üblichen Definition des Ranges einer Lieschen Gruppe vgl.: H. Hopf: *Über den Rang geschlossener Liescher Gruppen* [Comm. Math. Helv. 13 (1940/41)].

Ist nämlich $x = z_{i_1} \cdot z_{i_2} \cdot \cdots \cdot z_{i_r}$ ein beliebiges dieser Elemente, und ist $\bar{x} = z_{j_1} \cdot z_{j_2} \cdot \cdots \cdot z_{j_{l-r}}$ das Produkt derjenigen z_i, die in x nicht als Faktoren auftreten, so ist $\bar{x} \cdot x = \pm v_0$; ist aber y ein von $\bar{x}$ verschiedenes Element der Basis $[z_i]$ mit $d(y) = d(\bar{x})$, so ist $y \cdot x = 0$, weil y mindestens einen Faktor z_i enthält, der auch in x vorkommt.

In Nr. 2 sind wir ausgegangen von einem beliebigen irreduziblen Erzeugendensystem $(z_1, z_2, \cdots, z_l, 1)$ von $\mathfrak{R}(G)$; nach Nr. 3 sind den z_i die minimalen Elemente v_i zugeordnet, die eine Basis von $\mathfrak{B}(G)$ bilden. Man sieht ganz leicht, dass man auf diese Weise, d.h. durch geeignete Wahl der z_i, jede Basis von $\mathfrak{B}(G)$ erhalten kann (mit der unwesentlichen Einschränkung, dass man, wegen der Bedingung $z_1 \cdot z_2 \cdot \cdots \cdot z_l = v_0$, nur solche Basen von $\mathfrak{B}(G)$ bekommt, die durch eine lineare Transformation mit der Determinante 1 auseinander hervorgehen; das bedeutet nur, dass man eventuell v_1 durch ein rationales Vielfaches ersetzen muss.). Das beweist man z.B. mit Hilfe der Bemerkung, dass die Gleichungen $z_i \cdot v_k = \delta_{ik} v_0$, durch die bei gegebenen z_i die v_i vollständig bestimmt sind, invariant bleiben, wenn man auf die z_i und die v_i kontragrediente lineare Transformationen ausübt.

Man nehme die zu einem beliebigen irreduziblen Erzeugendensystem $(z_1, z_2, \cdots, z_l, 1)$ von $\mathfrak{R}(G)$ gehörenden minimalen Elemente v_i in eine sonst beliebige Basis von $\mathfrak{B}(G)$ auf. Dann ist das zu v_i duale Element z_i' der dualen Basis von der Form $z_i' = z_i + y_i$ mit nichtmaximalem y_i; das bestätigt man leicht auf Grund der Tatsache, dass $z_i \cdot v_i = v_0$ ist, dass aber für jedes von z_i und 1 verschiedene Element x der Basis $[z_i]$ der Schnitt $x \cdot v_i$ gleich Null ist. Diese Elemente z_i' bilden ebenfalls mit 1 zusammen ein irreduzibles Erzeugendensystem von $\mathfrak{R}(G)$; sie besitzen auch die Eigenschaften a) und b) (§1, Nr. 2); es ist nämlich $\delta(z_i') = \delta(z_i)$ und $z_1' \cdot z_2' \cdot \cdots \cdot z_l' = z_1 \cdot z_2 \cdot \cdots \cdot z_l$; aus den Gleichungen $z_i \cdot v_k = \delta_{ik} v_0$ folgen ohne weiteres die Gleichungen $z_i' \cdot v_k = \delta_{ik} v_0$.

3. Der Pontrjagin'sche Ring und der "Aufspann-Satz"

1. G sei, wie in §1, eine (geschlossene) Gruppenmannigfaltigkeit. Die Multiplikation in G, die jedem Punktepaar (p, q) von G den Punkt $p \cdot q$ zuordnet, kann man auffassen als eine Abbildung F des topologischen Produktes $G \times G$ zweier Exemplare von G in G, die definiert ist durch:

$$F(p \times q) = p \cdot q.$$

Sind x, y zwei Homologieklassen von G, so wollen wir unter dem Pontrjagin'schen Produkt [3] *(kürzer: P-Produkt) von x und y das Bild des Elementes $x \times y$ von $\mathfrak{R}(G \times G)$ bei F verstehen; wir bezeichnen es mit $x \circ y$; also:*

$$F(x \times y) = x \circ y.$$

Anschaulich, aber etwas unscharf, kann man das so sagen: Durchläuft der Punkt p den Zyklus x, und der Punkt q den Zyklus y, dann durchläuft der Punkt pq den Zyklus $x \circ y$.

Es ist $d(x \circ y) = d(x) + d(y)$.

Wegen des Assoziativgesetzes in G ist diese Multiplikation auch assoziativ; für drei Homologieelemente x_1, x_2, x_3 gilt:

$$(x_1 \circ x_2) \circ x_3 = x_1 \circ (x_2 \circ x_3).$$

Das P-Produkt ist mit der Addition distributiv verknüpft, wie man sofort sieht.

Das bedeutet: *Man kann die Betti'sche Gruppe $\mathfrak{B}(G)$ nicht nur durch die Schnittbildung $x \cdot y$ zu einem Ring $\mathfrak{R}(G)$, sondern auch durch die Pontrjagin'sche Produktbildung $x \circ y$ zu einem Ring $\mathfrak{P}(G)$, dem Pontrjagin'schen Ring von G, machen.*

Auf den Ring $\mathfrak{P}(G)$ kann man die üblichen algebraischen Begriffe anwenden; sind $x_1, x_2, \cdots x_r$ irgendwelche Elemente von $\mathfrak{P}(G)$, so ist klar, was man unter dem von den x_i erzeugten Pontrjagin'schen Teilring und was man unter dem von den x_i erzeugten Pontrjagin'schen Ideal (Rechts-, Links-, oder zweiseitig) zu verstehen hat.

(Jedes von homogen-dimensionalen x_i erzeugte Ideal in $\mathfrak{R}(G)$ ist übrigens zweiseitig, wegen der Regel $x \cdot y = \pm y \cdot x$ für homogen-dimensionale x, y; es wird sich zeigen, dass das auch in $\mathfrak{P}(G)$ so ist.)

Der Ring $\mathfrak{P}(G)$ besitzt eine Eins, nämlich das nulldimensionale Element v_0, das man etwa durch den einfach gezählten Einheitspunkt e, das Gruppen-Einselement von G, repräsentieren kann; denn wegen $F(e \times p) = ep = p$ für jeden Punkt p von G ist offenbar $F(v_0 \times x) = x$ für jedes Element x von $\mathfrak{P}(G)$, und ebenso ist $F(x \times v_0) = x$, d.h.:

$$v_0 \circ x = x \circ v_0 = x.$$

2. Der Ring $\mathfrak{P}(G)$ und sein Verhältnis zu dem Ring $\mathfrak{R}(G)$ soll nun untersucht werden.

Die Pontrjagin'schen Untersuchungen [3] haben ergeben, dass für die Gruppen der bekannten vier Killing-Cartan'schen Klassen die Ringe $\mathfrak{R}(G)$ und $\mathfrak{P}(G)$ ein ganz bestimmtes, in gewissem Sinn duales, Verhalten aufweisen. Von Cartan [15] und in anderer Form von Hopf [16] ist die Vermutung ausgesprochen worden, dass diese Dualität den Ringen $\mathfrak{R}(G)$ und $\mathfrak{P}(G)$ aller (geschlossener) Gruppenmannigfaltigkeiten zukomme. Diese Vermutung wollen wir im folgenden beweisen. Satz I ist die Cartan'sche, Satz I′ die Hopf'sche Formulierung der Vermutung. Aus Gründen, die bei der Formulierung und beim Beweis der Vermutung, besonders in der Form I′, klar werden werden, wollen wir die zu beweisende Tatsache den "Aufspann-Satz" nennen. Die Cartan'sche Formulierung ist [15]:

Satz I. *G sei eine geschlossene Gruppen-Mannigfaltigkeit der Dimension n, vom Rang l.*

a) *Dann gibt es im Ring $\mathfrak{P}(G)$ (mit rationalen Koeffizienten) l Elemente $v_1, v_2, \cdots v_l$ ungerader Dimensionen, für die gilt:*

$$v_i \circ v_k = -v_k \circ v_i \quad \text{und} \quad v_i \circ v_i = 0;$$

die $2^l - 1$ Elemente

$$v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r} \qquad (i_1 < i_2 < \cdots < i_r)$$

bilden mit dem Punkt v_0 zusammen eine volle Betti'sche Basis (Basis der Betti'schen Gruppe $\mathfrak{B}(G)$).

b) *Sind $w_1 = v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r}$ und $w_2 = v_{j_1} \circ v_{j_2} \circ \cdots \circ v_{j_s}$ zwei solche Basiselemente, so ist der Schnitt $w_1 \cdot w_2$ gleich Null, wenn eines der Elemente v_i weder in dem Produkt w_1 noch in dem Produkt w_2 als Faktor vorkommt; kommt jedes v_i entweder in w_1 oder in w_2 oder in beiden als Faktor vor, dann ist der Schnitt $w_1 \cdot w_2$ (bis auf das Vorzeichen) gleich dem Produkt $v_{k_1} \circ v_{k_2} \circ \cdots \circ v_{k_t}$ $(k_1 < k_2 < \cdots < k_t)$ derjenigen v_i, die sowohl in w_1 als auch in w_2 als Faktor auftreten, oder, falls kein v_i sowohl in w_1 als auch in w_2 vorkommt, gleich v_0.*

Teil a) von Satz I beschreibt den Ring $\mathfrak{P}(G)$; die Elemente v_i sind antikommutativ und bilden ein irreduzibles Erzeugendensystem von $\mathfrak{P}(G)$; aus dem Beweis wird hervorgehen, dass man für diese Elemente v_i die Elemente einer beliebigen Basis der Gruppe $\mathfrak{B}(G)$ der minimalen Elemente von $\mathfrak{R}(G)$ wählen darf. Die in I a) genannte Basis von $\mathfrak{B}(G)$ wollen wir die Basis $\langle v_i \rangle$ nennen; sie ist durch die Elemente v_i bestimmt.

Teil b) enthält die Beziehung zwischen $\mathfrak{R}(G)$ und $\mathfrak{P}(G)$; denn b) bestimmt den Schnitt zweier beliebiger Elemente von $\mathfrak{P}(G)$. Diese Beziehung tritt klarer hervor in der Hopf'schen Formulierung des Satzes I, die durch den Satz I' gegeben wird. Dazu ist eine Vorbereitung nötig.

3. $\Pi = S_{m_1} \times S_{m_2} \times S_{m_3} \times \cdots \times S_{m_l}$ sei ein Produkt von Sphären S_{m_i} der Dimensionen m_i. Z_i und V_i seien die in §2, Nr. 2 u. 3 definierten maximalen und minimalen Elemente von $\mathfrak{R}(\Pi)$.

In $\mathfrak{B}(\Pi)$ lässt sich neben der Schnittbildung in naheliegender Weise noch eine zweite Multiplikation, das "Aufspannen," erklären [16]. Das Produkt zweier Elemente X, Y von $\mathfrak{B}(\Pi)$ wird mit $X \boxtimes Y$ bezeichnet; die Definition geschieht folgendermassen:

Für $i_1 < i_2 < \cdots < i_r$ ist $V_{i_1} \boxtimes V_{i_2} \boxtimes \cdots \boxtimes V_{i_r}$ dasjenige Element $x_1 \times x_2 \times \cdots \times x_l$ von $\mathfrak{B}(\Pi)$, in dem

$$x_j = S_{m_j} \quad \text{für} \quad j = i_1, i_2, \cdots, i_r \text{ und}$$
$$x_j = p \qquad \text{für alle anderen } j$$

ist. Das sind mit dem Punkt V_0 zusammen 2^l Elemente, die eine volle Betti'sche Basis von $\mathfrak{B}(\Pi)$ bilden.

Für zwei solche Elemente

$$X = V_{i_1} \boxtimes V_{i_2} \boxtimes \cdots \boxtimes V_{i_r}, \qquad Y = V_{j_1} \boxtimes V_{j_2} \boxtimes \cdots \boxtimes V_{j_s}$$

wird das "aufgespannte" Element $X \boxtimes Y$ dadurch definiert, dass man die rechten Seiten nach dem assoziativen Gesetz miteinander multipliziert und die Regeln

$$V_i \boxtimes V_k = (-1)^{m_i \cdot m_k} V_k \boxtimes V_i, \qquad V_i \boxtimes V_i = 0$$

$(m_i = d(V_i))$ anwendet; dann wird $X \boxtimes Y$ entweder $= 0$ oder (bis auf das Vorzeichen) eins der oben genannten Basiselemente. Für beliebige Elemente

von $\mathfrak{B}(\Pi)$ definiert man das Produkt, indem man sie durch die Elemente dieser Basis darstellt und die distributiven Gesetze anwendet. Es ist $d(X \boxtimes Y) = d(X) + d(Y)$.

Durch diese Multiplikation wird die Betti'sche Gruppe $\mathfrak{B}(\Pi)$ zu einem Ring $\mathfrak{Q}(\Pi)$. Es existiert eine Eins:

$$V_0 = p \times p \times \cdots \times p.$$

Die Ringe $\mathfrak{R}(\Pi)$ und $\mathfrak{Q}(\Pi)$ sind isomorph; die Isomorphie bekommt man, wenn man die Erzeugenden Z_i von $\mathfrak{R}(\Pi)$ den Erzeugenden V_i von $\mathfrak{Q}(\Pi)$ zuordnet.

Es besteht aber weiter eine gewisse Dualität; es gilt nämlich die folgende "Formel (Δ)":

$$(\Delta) \qquad Z_{i_1} \cdot Z_{i_2} \cdot \cdots \cdot Z_{i_r} = \pm V_{j_1} \boxtimes V_{j_2} \boxtimes \cdots \boxtimes V_{j_{l-r}},$$

wenn die Mengen $\{i_1, i_2, \cdots, i_r\}$ und $\{j_1, j_2, \cdots, j_{l-r}\}$ komplementäre Teilmengen der Menge der Zahlen von 1 bis l sind. Beide Seiten bedeuten nämlich dasjenige Element $x_1 \times x_2 \times \cdots \times x_l$ von $\mathfrak{B}(\Pi)$, in dem

$$x_j = S_{m_j} \quad \text{für} \quad j = j_1, j_2, \cdots, j_{l-r} \text{ und}$$
$$x_j = p \quad \text{für} \quad j = i_1, i_2, \cdots, i_r$$

ist. Das Vorzeichen hängt auf gewisse Weise von den Indizes ab.

Mit Hilfe dieser Formel (Δ) lässt sich jedes Element von $\mathfrak{B}(\Pi)$ sowohl als Element von $\mathfrak{R}(\Pi)$ als auch als Element von $\mathfrak{Q}(\Pi)$ schreiben. Infolgedessen kann man den Schnitt $X \cdot Y$ zweier beliebiger Elemente X, Y von $\mathfrak{Q}(\Pi)$ und das Produkt $X \boxtimes Y$ zweier beliebiger Elemente X, Y von $\mathfrak{R}(\Pi)$ berechnen; die Formel (Δ) bestimmt die Beziehung zwischen $\mathfrak{R}(\Pi)$ und $\mathfrak{Q}(\Pi)$ vollkommen.

4. Der Satz I' lautet nun folgendermassen [16]:

SATZ I'. *Die auf Grund von Satz A mögliche, isomorphe Abbildung der Ringe* $\mathfrak{R}(G)$ *und* $\mathfrak{R}(\Pi)$ (*wo* Π *ein gewisses Sphärenprodukt ist*) *lässt sich so wählen, dass zugleich die Ringe* $\mathfrak{P}(G)$ *und* $\mathfrak{Q}(\Pi)$ *isomorph aufeinander abgebildet werden.*

Der Beweis für die Sätze I und I' wird in den Nummern 5–8 geliefert; er geht folgendermassen vor sich: Wir beweisen zunächst (in Nr. 5, 6, 7) den Teil a) von Satz I, und zwar in der folgenden verschärften Form: Die Elemente v_i einer beliebigen Basis $(v_1, v_2, \cdots v_l)$ der Gruppe $\mathfrak{B}(G)$ der minimalen Elemente von $\mathfrak{R}(G)$ haben die in I a) genannten Eigenschaften. Den Teil b) von Satz I beweisen wir in Nr. 8 folgendermassen: Wir bilden mit den Elementen v_i die in I a) genannte Basis $\langle v_i \rangle$; die in der dualen Basis zu den v_i dualen Elemente nennen wir z_i; sie haben die in §2, Nr. 4, Ende, genannten Eigenschaften. Dann beweisen wir die Gültigkeit der der Formel (Δ) (Nr. 3) entsprechenden "Formel (D)":

$$(D) \qquad z_{i_1} \cdot z_{i_2} \cdot \cdots \cdot z_{i_r} = \pm v_{j_1} \circ v_{j_2} \circ \cdots \circ v_{j_{l-r}},$$

für komplementäre Teilmengen $\{i_1, i_2, \cdots, i_r\}$ und $\{j_1, j_2, \cdots, j_{l-r}\}$ der Menge $\{1, 2, \cdots, l\}$.

Aus dieser Formel folgt dann sofort I b): Man schreibe die Elemente w_1 und w_2 mit Hilfe von (D) als Schnitt gewisser z_i ; dann bilde man den Schnitt $w_1 \cdot w_2$, der entweder 0 oder wieder ein Schnitt gewisser z_i ist; im zweiten Falle schreibt man $w_1 \cdot w_2$ wieder mit Hilfe von (D) als P-Produkt gewisser v_i . Man bestätigt ohne Mühe, dass der Schnitt $w_1 \cdot w_2$ das in I b) angegebene Verhalten zeigt, wegen der Dualität der Formel (D), die man folgendermassen aussprechen kann: Kommt in (D) auf der linken Seite das Element z_i vor, so kommt rechts das dazu duale v_i nicht vor; kommt links das Element z_k nicht vor, so kommt rechts das dazu duale v_k vor.

Der Satz I′ ist dann auch bewiesen: ordnet man die eben genannten z_i den Elementen Z_i von $\mathfrak{R}(\Pi)$ zu, so erhält man einen Isomorphismus $\mathfrak{R}(G) \approx \mathfrak{R}(\Pi)$. Dabei entsprechen wegen §2, Nr. 4 die Elemente v_i von $\mathfrak{R}(G)$ den Elementen V_i von $\mathfrak{R}(\Pi)$. Wegen der Formeln (D) und (Δ) entspricht dann dem P-Produkt irgendwelcher v_{i_k} das von den entsprechenden V_{i_k} "aufgespannte" Element, d. h. es liegt ein Isomorphismus zwischen $\mathfrak{P}(G)$ und $\mathfrak{Q}(\Pi)$ vor; um Uebereinstimmung der Vorzeichen in (D) und (Δ) zu erhalten, muss man jedoch die Elemente Z_i durch die Elemente $(-1)^{l-i} Z_i$ ersetzen (vgl. §2, Nr. 3).

5. Wir beginnen mit dem Beweis von I und I′. Wir betrachten die Abbildung

$$F(p \times q) = pq$$

des Produktes $G \times G$ zweier Exemplare von G in G; ϕ sei der zugehörige Umkehrungshomomorphismus (§1, Nr. 5).

Das Bild $\phi(z)$ eines beliebigen Elementes z von $\mathfrak{R}(G)$ hat eine sehr spezielle Gestalt, auf der alle weiteren Ueberlegungen beruhen; sie ist eine Folge der Tatsache, dass der Punkt v_0 in $\mathfrak{P}(G)$ die Rolle der Eins spielt, also eine Folge der Gleichungen

$$F(x \times v_0) = x, \qquad F(v_0 \times y) = y$$

für beliebige Elemente x, y von $\mathfrak{B}(G)$.

Das Bild $\phi(z)$ eines beliebigen $z(\neq 1)$ lässt sich in der folgenden Form schreiben:

$$\phi(z) = (1 \times y) + (x \times 1) + \sum (x_h \times y_h),$$

wo die Ungleichungen $d(z) < d(x_h) < n$, $d(z) < d(y_h) < n$ für alle h gelten; denn auf die Form der rechten Seite kann man jedes Element von $\mathfrak{R}(G \times G)$ bringen, und die Ungleichungen folgen aus der Gleichung $\delta(\phi(z)) = \delta(z)$. Hierin lassen sich nun y und x bestimmen:

Wir bilden den Schnitt $\phi(z) \cdot (v_0 \times 1)$. Der Schnitt von v_0 mit nicht skalaren Elementen ist immer 0; auf Grund der in §1, Nr. 7 genannten Sätze findet man sofort:

$$\phi(z) \cdot (v_0 \times 1) = v_0 \times y.$$

Uebt man auf diese Gleichung die Abbildung F aus und beachtet die Funktionalgleichung für ϕ (§1, Nr. 5), so erhält man:

$$z \cdot 1 = y, \quad \text{also} \quad z = y.$$

Entsprechend bilden wir den Schnitt $\phi(z) \cdot (1 \times v_0)$. Man findet:

$$\phi(z) \cdot (1 \times v_0) = (-1)^{\delta(z) \cdot \delta(v_0)} (x \times v_0).$$

Durch Ausübung von F und Anwendung der Funktionalgleichung für ϕ erhält man, wenn man noch beachtet, dass $\delta(v_0) = n$ ist:

$$z = (-1)^{n \cdot \delta(z)} x.$$

Damit ist gezeigt: Für jedes Element $z(\neq 1)$ von $\mathfrak{R}(G)$ gilt die "Formel (ϕ)":

$$(\phi) \qquad \phi(z) = (1 \times z) + (-1)^{n \cdot \delta(z)} (z \times 1) + \sum (x_h \times y_h)$$

mit den Ungleichungen $d(z) < d(x_h) < n$, $d(z) < d(y_h) < n$. Ausserdem ist natürlich $\phi(1) = 1 \times 1$.

Wir führen nun eine Basis in $\mathfrak{B}(G)$ ein. Und zwar wählen wir die von einem beliebigen irreduziblen Erzeugendensystem $(1, z_1, z_2, \cdots z_l)$ von $\mathfrak{R}(G)$ erzeugte Basis $[z_i]$ (§2, Nr. 2); wir dürfen $z_1 \cdot z_2 \cdot \cdots \cdot z_l = v_0$ annehmen. Das Vorzeichen der Elemente v_i, die zu dieser Basis gehören, ist durch

$$v_i = (-1)^{i-1} z_1 \cdot z_2 \cdot \cdots \cdot z_{i-1} \cdot z_{i+1} \cdot \cdots \cdot z_l$$

definiert; dann sind in der dualen Basis, die nach §2, Nr. 4 bis auf Vorzeichen aus denselben Elementen wie die Basis $[z_i]$ besteht, gerade die Elemente z_i zu den Elementen v_i dual, d. h. es ist $z_i \cdot v_i = v_0$; es gelten dann die Gleichungen:

$$z_i \cdot v_k = \delta_{ik} v_0 .$$

Die Elemente v_i bilden eine Basis von $\mathfrak{B}(G)$; nach §2, Nr. 4 kann man durch geeignete Wahl der Elemente z_i jede Basis von $\mathfrak{B}(G)$ auf diese Weise erhalten.

Durch die Basis $[z_i]$ in $\mathfrak{B}(G)$ ist eine Basis in $\mathfrak{B}(G \times G)$ bestimmt (§1, Nr. 7), die wir die Basis $[z_i]^\times$ nennen wollen; auch diese Basis ist (bis auf Vorzeichen) zu sich dual.

Für die maximalen Elemente z_i ist $\delta(z_i)$ ungerade. Daher heisst für sie die Formel (ϕ):

$$(\phi_i) \qquad \phi(z_i) = (1 \times z_i) + (-1)^n (z_i \times 1) + R_i \, ;$$

dabei ist der Term $\sum (x_h \times y_h)$ aus (ϕ) mit R_i bezeichnet worden. Weil alle $\delta(z_i)$ ungerade sind, ist für jeden Term $(x_h \times y_h)$ von R_i entweder $\delta(x_h)$ oder $\delta(y_h)$ gerade, also nie sowohl x_h als auch y_h maximal, sondern wenigstens eines zusammengesetzt.

6. Wir beweisen zunächst die Behauptungen

$$v_i \circ v_i = 0 \quad \text{und} \quad v_i \circ v_k = -v_k \circ v_i$$

von I a); anders gesagt, wir beweisen: für die Elemente $v_i \times v_i$ und $(v_i \times v_k) + (v_k \times v_i)$, die wir für den Moment mit v_{ii} und v_{ik} bezeichnen, gilt:

$$F(v_{ii}) = F(v_{ik}) = 0.$$

Zum Beweis beachten wir, dass die Funktionalgleichung für ϕ die Gleichung $F(\phi(x')\cdot v_{ii}) = x'\cdot F(v_{ii})$ und die entsprechende für v_{ik} mit einem beliebigen $x' \in \mathfrak{B}(G)$ liefert. Um zu zeigen, dass $F(v_{ii}) = 0$ ist, genügt es wegen des Poincaré-Veblen'schen Dualitätssatzes, die folgende Behauptung (a) zu beweisen:

(a) für jedes x von $\mathfrak{B}(G)$ mit $\delta(x) = d(v_{ii})$ ist die Schnittzahl $\phi(x)\cdot v_{ii}$ gleich Null.

Ebenso genügt es für den Beweis von $F(v_{ik}) = 0$, zu zeigen:

(b) für jedes y von $\mathfrak{B}(G)$ mit $\delta(y) = d(v_{ik})$ ist die Schnittzahl $\phi(y)\cdot v_{ik}$ gleich Null.

Wir betrachten nun die Menge derjenigen Elemente $\sum (x_h \times y_h)$ in $\mathfrak{R}(G \times G)$, in denen in jedem Term $x_h \times y_h$ entweder x_h oder y_h oder beide zusammengesetzte Elemente von $\mathfrak{R}(G)$ sind; wie man sofort einsieht, ist diese Menge, die $\mathfrak{U}$ heisse, ein zweiseitiges Ideal in $\mathfrak{R}(G \times G)$. Die Elemente $(z_i \times z_i)$, $(z_i \times z_k)$, $(z_k \times z_i)$ gehören zu verschiedenen Restklassen von $\mathfrak{U}$.

Aus der Formel (ϕ_i) und der anschliessenden Bemerkung über R_i folgen die Kongruenzen:

$$\phi(z_r) \equiv (1 \times z_r) + (-1)^n (z_r \times 1) \qquad \text{mod } \mathfrak{U}.$$

Weil die Elemente $(1 \times z_r\cdot z_s)$ und $(z_r\cdot z_s \times 1)$ zu $\mathfrak{U}$ gehören, folgt daraus sofort:

$$\phi(z_r\cdot z_s) = \phi(z_r)\cdot\phi(z_s) \equiv (-1)^n (z_s \times z_r) - (-1)^n (z_r \times z_s) \qquad \text{mod } \mathfrak{U},$$

wo das Vorzeichen des zweiten Gliedes daraus folgt, dass die $\delta(z_j)$ ungerade sind. Ebenso rechnet man sofort aus:

$$\phi(z_r\cdot z_s\cdot z_t) = \phi(z_r)\cdot\phi(z_s)\cdot\phi(z_t) \equiv 0 \qquad \text{mod } \mathfrak{U},$$

und das Gleiche für alle Produkte mit mehr als drei Faktoren. Stellt man also für irgendein Element x der Basis $[z_i]$ von $\mathfrak{B}(G)$ das Bild $\phi(x)$ durch die Basis $[z_i]^\times$ von $\mathfrak{B}(G \times G)$ dar, so tritt darin das Basiselement $z_i \times z_i$ nicht (d. h. mit dem Koeffizienten Null) auf; die Basiselemente $z_i \times z_k$ und $z_k \times z_i$ treten nur dann mit nichtverschwindendem Koeffizienten auf, wenn $x = z_i\cdot z_k$ ist, und dann treten sie mit entgegensetzten Koeffizienten auf (nämlich mit $+1$ und -1).

Nun sind aber die Elemente $(z_i \times z_i)$, $(z_i \times z_k)$, $(z_k \times z_i)$, wenn man sie noch mit den Vorzeichen $(-1)^n$ versieht, die zu den Elementen $(v_i \times v_i)$, $(v_i \times v_k)$, $(v_k \times v_i)$ dualen Elemente der Basis $[z_i]^\times$ (man beachte, dass Basis und duale Basis bis auf Vorzeichen mit der Basis $[z_i]^\times$ übereinstimmen).

Daraus schliesst man sofort die Behauptung (a):

$$\phi(x)\cdot(v_i \times v_i) = 0, \quad \text{wenn} \quad \delta(x) = d(v_i \times v_i).$$

Auch für die Behauptung (b) muss man nur noch den Schnitt $\phi(z_i\cdot z_k)\cdot v_{ik}$ prüfen. Aber wegen der Verschiedenheit der Vorzeichen von $(z_i \times z_k)$ und $(z_k \times z_i)$ in $\phi(z_i\cdot z_k)$ ist auch dieser Schnitt Null. Damit ist (a) und (b), also auch die Behauptung am Anfang dieser Nummer, bewiesen.

7. Als zweites beweisen wir die Gleichung:

$$v_1 \circ v_2 \circ \cdots \circ v_l = 1.$$

Wegen der eben bewiesenen Antikommutativität der v_i dürfen wir annehmen, die v_i seien so angeordnet, dass $d(v_i) \leqq d(v_k)$ ist für $i < k$.

Dass das Produkt der v_i gerade $+1$ ist, liegt daran, dass die z_i so definiert sind, dass $z_1 \cdot z_2 \cdots z_l = +v_0$ ist; das Wesentliche ist, dass das—n-dimensionale—Element $v_1 \circ v_2 \circ \cdots \circ v_l$ nicht Null ist, d.h. dass die ganze Mannigfaltigkeit G von der Basis $\{v_1, v_2, \cdots v_l\}$ der Gruppe $\mathfrak{B}(G)$ "aufgespannt" wird; dem entspricht die Gleichung $V_1 \boxtimes V_2 \boxtimes \cdots \boxtimes V_l = 1$ beim Sphärenprodukt II.

Wir bezeichnen mit w_k das Teilprodukt $v_k \circ v_{k+1} \circ \cdots \circ v_l$ $(k = 1, 2, \cdots, l)$ und setzen noch $w_{l+1} = v_0$.

Dann beweisen wir für $k = 1, 2, \cdots, l$ die Gleichungen:

$$(1_k) \qquad z_k \cdot w_k = (-1)^{l-k} w_{k+1}$$

$$(2_k) \qquad z_i \cdot w_k = 0 \qquad \text{für } i < k.$$

Das ist richtig für $k = l$, wegen $z_l \cdot v_l = v_0$. Die Gleichungen seien bewiesen für k; wir beweisen sie für $k - 1$.

Wir gehen aus von der Tatsache

$$F(v_{k-1} \times w_k) = w_{k-1},$$

die aus der Definition der w_k folgt.

Für ein beliebiges Element z_i ergibt die Funktionalgleichung für ϕ die Gleichung:

$$(3) \qquad F(\phi(z_i) \cdot (v_{k-1} \times w_k)) = z_i \cdot w_{k-1}.$$

Für $i < k$ setze man jetzt $\phi(z_i)$ aus der Formel (ϕ_i) (Nr. 5) ein. Man erhält:

$$(4) \qquad \phi(z_i) \cdot (v_{k-1} \times w_k) = (-1)^{n+k-1}((z_i \cdot v_{k-1}) \times w_k) \qquad \text{für } i < k.$$

Denn wegen (2_k) ist $(1 \times z_i) \cdot (v_{k-1} \times w_k) = 0$; und ist $x \times y$ ein Element aus $R_i = \sum x_h \times y_h$, so ist wegen $d(x) > d(z_i) \geqq d(z_{k-1})$ der Schnitt $x \cdot v_{k-1} = 0$, also auch $R_i \cdot (v_{k-1} \times w_k) = 0$ (die zweite Ungleichung folgt aus der Voraussetzung über die $d(v_i)$).

Für $i = k - 1$ erhält man also:

$$(1') \qquad \phi(z_{k-1}) \cdot (v_{k-1} \times w_k) = (-1)^{l-(k-1)}(v_0 \times w_k),$$

für $i < k - 1$ erhält man dagegen:

$$(2') \qquad \phi(z_i) \cdot (v_{k-1} \times w_k) = 0 \qquad \text{für } i < k - 1.$$

Aus (1') und (3) entnimmt man jetzt:

$$z_{k-1} \cdot w_k = (-1)^{l-(k-1)} w_k,$$

aus (2') und (3) dagegen:

$$z_i \cdot w_{k-1} = 0 \qquad \text{für } i < k - 1;$$

und das sind die Gleichungen (1_{k-1}) und (2_{k-1}).

Aus den damit für $k = 1, 2, \cdots, l$ bewiesenen Gleichungen (1_k) schliesst man jetzt der Reihe nach: $w_l \neq 0$ (wegen $w_{l+1} = v_l \neq 0$), $w_{l-1} \neq 0, \cdots, w_1 \neq 0$. Also ist jedenfalls $w_1 = v_1 \circ v_2 \circ \cdots \circ v_l \neq 0$. Bei Durchführung der elementaren Rechnung erhält man die genauere Aussage:

$$v_1 \circ v_2 \circ \cdots \circ v_l = 1.$$

Aus den Tatsachen, dass die v_i antikommutativ sind (in $\mathfrak{P}(G)$), und dass $v_1 \circ v_2 \circ \cdots \circ v_l \neq 0$ ist, kann man nun schliessen, dass die Elemente v_i (zusammen mit v_0) den Ring $\mathfrak{P}(G)$ erzeugen. Dazu betrachten wir die 2^l Elemente

$$v_0 \quad \text{und} \quad v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r} \qquad (i_1 < i_2 < \cdots < i_r).$$

Wegen der Antikommutativität der v_i ist jedes beliebige Potenz-P-Produkt, das man aus den v_i bilden kann, entweder 0 oder—bis aufs Vorzeichen—einem dieser Elemente gleich. Nun stellt man leicht fest, dass diese Elemente linear unabhängig sind:

$0 = \sum a_i \cdot y_i$ sei eine Relation zwischen solchen Elementen y_i mit Koeffizienten a_i. Es sei y_1 ein "kürzestes" dieser Elemente, d. h. eines mit möglichst kleiner Zahl von Faktoren v_i; mit $\bar{y}_1$ werde das P-Produkt derjenigen v_i bezeichnet, die in y_1 nicht als Faktoren auftreten.

Nun multipliziere man die Gleichung $0 = \sum a_i \cdot y_i$ im Sinne des P-Produktes mit $\bar{y}_1$. Auf der rechten Seite ergibt $a_1 \cdot y_1 \circ \bar{y}_1$ das Element $\pm a_1 \cdot 1$; denn es ist $y_1 \circ \bar{y}_1 = \pm v_1 \circ v_2 \circ \cdots \circ v_l$. Für die y_i mit $i > 1$ ist aber $y_i \circ \bar{y}_1 = 0$, weil mindestens ein v_k in y_i und in $\bar{y}_1$ zugleich vorkommt. Man erhält also: $0 = \pm a_1$; ebenso zeigt man, dass alle a_i verschwinden.

Nun hat aber $\mathfrak{B}(G)$ den Rang 2^l (§2, Nr. 2). Also bilden die Elemente

$$v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r} \qquad (i_1 < i_2 < \cdots < i_r)$$

mit v_0 zusammen eine Basis von $\mathfrak{B}(G)$, die die Basis $\langle v_i \rangle$ genannt werden soll.

Da nach §2, Nr. 3 die $d(v_i)$ ungerade sind, ist jetzt I a) vollständig bewiesen. Die Elemente v_i bilden ein irreduzibles Erzeugendensystem von $\mathfrak{P}(G)$.

8. Wir kommen jetzt zum Beweis von I b). Wir betrachten wieder die Abbildungen F und ϕ. Aber als Basis in G und zur Bildung der Basis in $G \times G$ nehmen wir jetzt die Basis $\langle v_i \rangle$. In der dualen Basis sind zu den v_i dual gewisse Elemente z_i', die nach §2, Nr. 4 ein irreduzibles Erzeugendensystem von $\mathfrak{R}(G)$ bilden, und für die die Gleichungen

$$z_i' \cdot v_k = \delta_{ik} v_0$$

gelten. Die z_i' nennen wir wieder z_i; für die so durch die v_i bestimmten z_i beweisen wir die Formel (D).

Wir behaupten: für die so bestimmten z_i verschwindet der Rest R_i in der Formel (ϕ_i), d. h. es gilt die verfeinerte Formel (ϕ_0):

$$(\phi_0) \qquad \phi(z_i) = (1 \times z_i) + (-1)^n (z_i \times 1).$$

Zum Beweis betrachten wir die Abbildung F; sie ist vollständig zu übersehen. Die Bilder der Elemente der durch die Basis $\langle v_i\rangle$ von $\mathfrak{B}(G)$ bestimmten Basis $\langle v_i\rangle^\times$ von $\mathfrak{B}(G \times G)$ werden durch folgende Gleichungen gegeben:

$$F((v_{i_1}\circ v_{i_2}\circ \cdots \circ v_{i_r}) \times (v_{j_1}\circ v_{j_2}\circ \cdots \circ v_{j_s})) = v_{i_1}\circ v_{i_2}\circ \cdots \circ v_{i_r}\circ v_{j_1}\circ v_{j_2}\circ \cdots \circ v_{j_s}.$$

Dabei ist das Element rechts entweder 0 oder—bis aufs Vorzeichen—ein Element der Basis $\langle v_i\rangle$. ($v_{i_1}\circ v_{i_2}\circ \cdots \circ v_{i_r}$ oder $v_{j_1}\circ v_{j_2}\circ \cdots \circ v_{j_s}$ kann auch das Element v_0 bedeuten; dann ist F durch $F(v_0 \times x) = F(x \times v_0) = x$ gegeben.)

Nun seien w_1 und w_2 zwei beliebige Elemente der Basis $\langle v_i\rangle$ von $\mathfrak{B}(G)$, die der Bedingung $\delta(z_i) = d(w_1) + d(w_2)$ genügen, sodass also $w = w_1 \times w_2$ ein Element der Basis $\langle v_i\rangle^\times$ von $\mathfrak{B}(G \times G)$ mit $\delta(\phi(z_i)) = d(w)$ ist. Wir betrachten die Schnittzahl $\phi(z_i)\cdot w$. Wegen der Funktionalgleichung für ϕ ist

$$F(\phi(z_i)\cdot w) = z_i \cdot F(w).$$

Weil z_i das zu v_i duale Element ist, ist diese Schnittzahl nur dann nicht Null, wenn $F(w)$, das ja ein Element der Basis $\langle v_i\rangle$ von $\mathfrak{B}(G)$ ist, gleich v_i ist. Daraus folgt: Der Schnitt $\phi(z_i)\cdot w$ ist nur dann nicht Null, wenn entweder $w = v_i \times v_0$ oder $w = v_0 \times v_i$ ist. Das bedeutet: in der Darstellung von $\phi(z_i)$ durch die zu der Basis $\langle v_i\rangle^\times$ duale Basis treten nur die zu $v_i \times v_0$ und $v_0 \times v_i$ dualen Elemente mit von Null verschiedenem Koeffizienten auf; und das sind eben die Elemente $(1 \times z_i)$ und $(-1)^n(z_i \times 1)$. Damit ist die Formel (ϕ_0) bewiesen.

Wir betrachten ein beliebiges Produkt $v_{i_1}\circ v_{i_2}\circ \cdots \circ v_{i_r}$ (mit lauter verschiedenen Faktoren v). Wir behaupten das Bestehen der folgenden Gleichungen:

$$(1)\quad z_{i_1}\cdot(v_{i_1}\circ v_{i_2}\circ \cdots \circ v_{i_r}) = \begin{cases} v_0 & \text{für } r = 1, \\ (-1)^{r-1} v_{i_2}\circ v_{i_3}\circ \cdots \circ v_{i_r} & \text{für } r > 1, \end{cases}$$

(2) $z_k\cdot(v_{i_1}\circ v_{i_2}\circ \cdots \circ v_{i_r}) = 0$, wenn keiner der Indizes $i_1, i_2, \cdots, i_r$ gleich k ist

Beachtet man die Antikommutativität der v_i, so erhält man aus der Gleichung (1) die folgende etwas allgemeinere, in der i_k ein beliebiger der Indizes $i_1, i_2, \cdots, i_r$ ist:

$$(1')\quad z_{i_k}\cdot(v_{i_1}\circ v_{i_2}\circ \cdots \circ v_{i_k}\circ \cdots \circ v_{i_r}) = \pm v_{i_1}\circ v_{i_2}\circ \cdots \circ v_{i_{k-1}}\circ v_{i_{k+1}}\circ \cdots \circ v_{i_r} \quad (\text{für } r > 1).$$

Diese Gleichungen besagen: Der Schnitt eines Elementes $w = v_{i_1}\circ v_{i_2}\circ \cdots \circ v_{i_r}$ mit einem Element z_j ist dann und nur dann nicht Null, wenn das zu z_j duale v_j als Faktor in w auftritt; tritt v_j als Faktor in w auf, dann erhält man den Schnitt $z_j\cdot w$ (bis aufs Vorzeichen), indem man den Faktor v_j aus dem Produkt $v_{i_1}\circ v_{i_2}\circ \cdots \circ v_{i_r}$ weglässt.

Den Beweis führen wir durch Induktion nach r; für $r = 1$ ist die Richtigkeit von (1) und (2) klar. (1) und (2) seien also richtig für alle P-Produkte mit weniger als r Faktoren v_i.

Wir gehen aus von der Tatsache:

$$F(v_{i_1} \times (v_{i_2} \circ v_{i_3} \circ \cdots \circ v_{i_r})) = v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r},$$

die aus der Definition des P-Produktes folgt.

Nach der Funktionalgleichung für ϕ ist:

$$F(\phi(z_j) \cdot (v_{i_1} \times (v_{i_2} \circ v_{i_3} \circ \cdots \circ v_{i_r}))) = z_j \cdot (v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r}). \tag{3}$$

In dem darin auftretenden Produkt

$$\phi(z_j) \cdot (v_{i_1} \times (v_{i_2} \circ v_{i_3} \circ \cdots \circ v_{i_r})),$$

das wir für den Moment X nennen wollen, setzen wir für $\phi(z_j)$ den Ausdruck aus Formel (ϕ_0) ein. Wir erhalten:

$$X = (v_{i_1} \times (z_j \cdot (v_{i_2} \circ v_{i_3} \circ \cdots \circ v_{i_r}))) + (-1)^{r-1}((z_j \cdot v_{i_1}) \times (v_{i_2} \circ v_{i_3} \circ \cdots \circ v_{i_r})).$$

Ist erstens $j = i_1$, so ist nach Induktionsvoraussetzung $z_j \cdot (v_{i_2} \circ v_{i_3} \circ \cdots \circ v_{i_r})$ gleich Null; andererseits ist dann $z_j \cdot v_{i_1} = v_0$. Also ist

$$X = (-1)^{r-1} v_0 \times (v_{i_2} \circ v_{i_3} \circ \cdots \circ v_{i_r})$$

und die Gleichung (3) liefert die Gleichung (1) für den Index r.

Ist zweitens j keiner der Indizes $i_1, i_2, \cdots, i_r$, so ist nach Induktionsvoraussetzung $z_j \cdot (v_{i_2} \circ v_{i_3} \circ \cdots \circ v_{i_r})$ gleich Null. Aber auch $z_j \cdot v_{i_1}$ ist gleich Null; also ist $X = 0$. Dann liefert die Gleichung (3) die Gleichung (2) für den Index r.

Durch die Gleichungen (1) und (2) ist das gegenseitige Verhalten von $\mathfrak{R}(G)$ und $\mathfrak{P}(G)$ vollkommen bestimmt; wir wollen jetzt die Dualitätsformel (D) daraus ableiten:

$$z_{i_1} \cdot z_{i_2} \cdots z_{i_r} = \pm v_{j_1} \circ v_{j_2} \circ \cdots \circ v_{j_{l-r}}, \tag{D}$$

wenn $\{i_1, i_2, \cdots, i_r\}$ und $\{j_1, j_2, \cdots, j_{l-r}\}$ komplementäre Teilmengen der Menge der Zahlen von 1 bis l sind.

(Das Vorzeichen hängt auf etwas komplizierte Weise von den Indizes ab; es sei bemerkt, dass

$$z_1 \cdot z_2 \cdots z_r = (-1)^{r(l-r)} v_{r+1} \circ v_{r+2} \circ \cdots \circ v_l$$

ist. Dass die Vorzeichen in (D) und (Δ) übereinstimmen (nach der Ersetzung von Z_i durch $(-1)^{l-i} Z_i$, vgl. Nr. 4), ergibt sich nach einer elementaren Rechnung.)

Wir gehen aus von der Tatsache

$$v_1 \circ v_2 \circ \cdots \circ v_l = 1,$$

und betrachten ein beliebiges Produkt $z_{i_1} \cdot z_{i_2} \cdot \cdots \cdot z_{i_r}$. Wir multiplizieren die Gleichung $v_1 \circ v_2 \circ \cdots \circ v_l = 1$ mit z_{i_1}. Rechts entsteht dabei z_{i_1}; und links fällt nach Gleichung (1′) der Faktor v_{i_1} weg. Die so entstandene Gleichung multiplizieren wir mit z_{i_2}. Rechts entsteht dabei $z_{i_1} \cdot z_{i_2}$; und links fällt auch noch

der Faktor v_{i_2} weg. So fahren wir fort; jedesmal kommt rechts ein Faktor z_{i_k} (im Schnitt-Sinn) dazu, und links fällt der P-Faktor v_{i_k} weg. So entsteht die Gleichung (D). Nach Nr. 5 ist damit I b), und damit der ganze "Aufspann-Satz" bewiesen.

9. Zum Rechnen in $\mathfrak{B}(G)$ bemerken wir noch folgendes:

Gehen die Elemente $v'_{i_1}, v'_{i_2}, \cdots, v'_{i_r}$ aus den Elementen $v_{i_1}, v_{i_2}, \cdots, v_{i_r}$ durch eine lineare Transformation

$$v'_{i_j} = \sum a_{jk} v_{i_k} \qquad (j, k = 1, 2, \cdots, r)$$

mit der Determinante $|a_{jk}|$ hervor, so gilt wegen der Antikommutativität der v_i die Gleichung:

$$v'_{i_1} \circ v'_{i_2} \circ \cdots \circ v'_{i_r} = |a_{jk}| v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r}.$$

Daraus folgt z. B., dass das Produkt (in $\mathfrak{B}(G)$) von linear abhängigen minimalen Elementen $v'_{i_1}, v'_{i_2}, \cdots, v'_{i_r}$ immer Null ist.

Weiter gilt: für beliebige (homogen-dimensionale) Elemente x, y ist

$$x \circ y = (-1)^{d(x) \cdot d(y)} y \circ x;$$

zum Beweis muss man nur x und y in der Basis $\langle v_i \rangle$ schreiben.

In Nr. 5 sind wir ausgegangen von einer beliebigen Basis $[z_i]$ von $\mathfrak{B}(G)$; dadurch wurden die Elemente v_i, die eine Basis von $\mathfrak{B}(G)$ bilden, bestimmt; nach §2, Nr. 4 kann man jede Basis von $\mathfrak{B}(G)$ auf diese Weise bekommen. In der Fassung I des Aufspannsatzes kann man also unter den Elementen v_i die Elemente einer beliebigen Basis von $\mathfrak{B}(G)$ verstehen. Die Elemente z_i, die dann in der Formel (D) auftreten (und die ein irreduzibles Erzeugendensystem von $\mathfrak{R}(G)$ bilden), sind folgendermassen bestimmt: sie sind in der zu der Basis $\langle v_i \rangle$ dualen Basis die zu den v_i dualen Elemente. Auf Grund der Formel (D) bestehen die Basis $[z_i]$ und die Basis $\langle v_i \rangle$ aus denselben Elementen; nun ist die Basis $[z_i]$ zu sich selbst dual (bis auf Vorzeichen); also ist die Basis $[z_i]$ zu der Basis $\langle v_i \rangle$ dual. Man bestätigt auch leicht unmittelbar an der Formel (D), dass das einzige Element der Basis $[z_i]$, das mit dem Element $v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r}$ der Basis $\langle v_i \rangle$ den Schnitt $\pm v_0$ ergibt, das Element $z_{i_1} \cdot z_{i_2} \cdot \cdots \cdot z_{i_r}$ ist. Wir wollen diese und noch zwei andere Regeln notieren, die unmittelbar aus (D) folgen:

(1) $$z_{i_1} \cdot z_{i_2} \cdot \cdots \cdot z_{i_r} \cdot (v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r}) = \pm v_0,$$

(2) $$z_{i_1} \cdot z_{i_2} \cdot \cdots \cdot z_{i_r} \cdot (v_{j_1} \circ v_{j_2} \circ \cdots \circ v_{j_r}) = 0, \qquad \text{wenn}$$

die Indexmengen $\{i_1, i_2, \cdots, i_r\}$ und $\{j_1, j_2, \cdots, j_r\}$ voneinander verschieden sind,

(3) $$z_{i_1} \cdot z_{i_2} \cdot \cdots \cdot z_{i_r} \cdot (v_{k_1} \circ v_{k_2} \circ \cdots \circ v_{k_s}) = 0, \qquad \text{wenn } s < r.$$

Um sie zu beweisen, hat man nur die P-Produkte der v_i mit Hilfe von (D) als Schnitt der z_i zu schreiben, und dann die Rechenregeln für $\mathfrak{R}(G)$ anzuwenden.

10. Es seien $x = z_{i_1} \cdot z_{i_2} \cdot \cdots \cdot z_{i_r}$ und $y = z_{j_1} \cdot z_{j_2} \cdot \cdots \cdot z_{j_{l-r}}$ zwei Elemente der Basis $[z_i]$, und es sei $x \cdot y = \pm v_0$, d. h. die Elemente $z_{i_1}, \cdots, z_{i_r}, z_{j_1}, \cdots, z_{j_{l-r}}$ sind die Elemente $z_1, z_2, \cdots, z_l$, nur in anderer Reihenfolge. Auf Grund von (D) ist dann

$$x \circ y = v_{j_1} \circ v_{j_2} \circ \cdots \circ v_{j_{l-r}} \circ v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r};$$

und dieses Produkt ist gleich ± 1, weil die darin auftretenden v_i bis auf die Reihenfolge mit den Elementen $v_1, v_2, \cdots, v_l$ übereinstimmen. Führt man die elementare Bestimmung des Vorzeichens durch, so erhält man: Aus $x \cdot y = v_0$ folgt $x \circ y = 1$. Daraus folgt leicht allgemein: Sind X, Y zwei beliebige Elemente von $\mathfrak{B}(G)$ mit $d(X) + d(Y) = n$, dann ist die Schnittzahl von X und Y gleich ihrer "Aufspannzahl," d. h. in $X \cdot Y = a \cdot v_0$ und in $X \circ Y = b \cdot 1$ sind die Koeffizienten a und b gleich; diese Gleichheit ist ja anschaulich sehr naheliegend.

Eine letzte Bemerkung ist folgende:

Die Elemente $v_1, \cdots, v_r$ seien irgendwelche, linear unabhängige Elemente von $\mathfrak{B}(G)$; dann sind die Elemente

$$v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_s} \qquad (1 \leqq i_k \leqq r,\ i_1 < i_2 < \cdots < i_s)$$

linear unabhängige Elemente von $\mathfrak{B}(G)$. Zum Beweis hat man nur die $v_1, \cdots, v_r$ in eine Basis $v_1, \cdots, v_l$ von $\mathfrak{B}(G)$ aufzunehmen, und mit dieser Basis die Basis $\langle v_i \rangle$ von $\mathfrak{B}(G)$ zu bilden.

KAPITEL II

ERSTE ANWENDUNGEN DES AUFSPANNSATZES

Der Aufspannsatz soll jetzt mit den wichtigsten gruppentheoretischen Begriffen in Zusammenhang gebracht werden: mit dem der homomorphen Abbildung und hauptsächlich denen der Untergruppe und der Nebengruppenzerlegung.

1. Homomorphe Abbildung

1. G und G' seien zwei Gruppenmannigfaltigkeiten. Eine Abbildung h von G in G' heisst homomorphe Abbildung, wenn sie

a) eine homomorphe Abbildung im gewöhnlichen gruppentheoretischen Sinn ist, und

b) eine stetige Abbildung der Mannigfaltigkeit G in die Mannigfaltigkeit G' ist.

Die homomorphe Abbildung h bildet den Ring $\mathfrak{P}(G)$ homomorph in den Ring $\mathfrak{P}(G')$ ab; die additive Homomorphie ist klar; die Gleichung

$$h(x \circ y) = h(x) \circ h(y)$$

für zwei beliebige Homologieklassen x, y von G und ihre Bilder $h(x)$, $h(y)$ bei h beweist man etwa folgendermassen:

Man bilde das topologische Produkt $G \times G$ durch die Abbildung $f(p \times q) = h(pq)$ in G' ab. Ist F die Abbildung $F(p \times q) = pq$ von $G \times G$ in G, ist F' die

der Faktor v_{i_2} weg. So fahren wir fort; jedesmal kommt rechts ein Faktor z_{i_k} (im Schnitt-Sinn) dazu, und links fällt der P-Faktor v_{i_k} weg. So entsteht die Gleichung (D). Nach Nr. 5 ist damit I b), und damit der ganze "Aufspann-Satz" bewiesen.

9. Zum Rechnen in $\mathfrak{P}(G)$ bemerken wir noch folgendes:

Gehen die Elemente $v'_{i_1}, v'_{i_2}, \cdots, v'_{i_r}$ aus den Elementen $v_{i_1}, v_{i_2}, \cdots, v_{i_r}$ durch eine lineare Transformation

$$v'_{i_j} = \sum a_{jk} v_{i_k} \qquad (j, k = 1, 2, \cdots, r)$$

mit der Determinante $|a_{jk}|$ hervor, so gilt wegen der Antikommutativität der v_i die Gleichung:

$$v'_{i_1} \circ v'_{i_2} \circ \cdots \circ v'_{i_r} = |a_{jk}| v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r}.$$

Daraus folgt z. B., dass das Produkt (in $\mathfrak{P}(G)$) von linear abhängigen minimalen Elementen $v'_{i_1}, v'_{i_2}, \cdots, v'_{i_r}$ immer Null ist.

Weiter gilt: für beliebige (homogen-dimensionale) Elemente x, y ist

$$x \circ y = (-1)^{d(x) \cdot d(y)} y \circ x;$$

zum Beweis muss man nur x und y in der Basis $\langle v_i \rangle$ schreiben.

In Nr. 5 sind wir ausgegangen von einer beliebigen Basis $[z_i]$ von $\mathfrak{B}(G)$; dadurch wurden die Elemente v_i, die eine Basis von $\mathfrak{B}(G)$ bilden, bestimmt; nach §2, Nr. 4 kann man jede Basis von $\mathfrak{B}(G)$ auf diese Weise bekommen. In der Fassung I des Aufspannsatzes kann man also unter den Elementen v_i die Elemente einer beliebigen Basis von $\mathfrak{B}(G)$ verstehen. Die Elemente z_i, die dann in der Formel (D) auftreten (und die ein irreduzibles Erzeugendensystem von $\mathfrak{R}(G)$ bilden), sind folgendermassen bestimmt: sie sind in der zu der Basis $\langle v_i \rangle$ dualen Basis die zu den v_i dualen Elemente. Auf Grund der Formel (D) bestehen die Basis $[z_i]$ und die Basis $\langle v_i \rangle$ aus denselben Elementen; nun ist die Basis $[z_i]$ zu sich selbst dual (bis auf Vorzeichen); also ist die Basis $[z_i]$ zu der Basis $\langle v_i \rangle$ dual. Man bestätigt auch leicht unmittelbar an der Formel (D), dass das einzige Element der Basis $[z_i]$, das mit dem Element $v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r}$ der Basis $\langle v_i \rangle$ den Schnitt $\pm v_0$ ergibt, das Element $z_{i_1} \cdot z_{i_2} \cdot \cdots \cdot z_{i_r}$ ist. Wir wollen diese und noch zwei andere Regeln notieren, die unmittelbar aus (D) folgen:

(1) $$z_{i_1} \cdot z_{i_2} \cdot \cdots \cdot z_{i_r} \cdot (v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r}) = \pm v_0,$$

(2) $$z_{i_1} \cdot z_{i_2} \cdot \cdots \cdot z_{i_r} \cdot (v_{j_1} \circ v_{j_2} \circ \cdots \circ v_{j_r}) = 0, \qquad \text{wenn}$$

die Indexmengen $\{i_1, i_2, \cdots, i_r\}$ und $\{j_1, j_2, \cdots, j_r\}$ voneinander verschieden sind,

(3) $$z_{i_1} \cdot z_{i_2} \cdot \cdots \cdot z_{i_r} \cdot (v_{k_1} \circ v_{k_2} \circ \cdots \circ v_{k_s}) = 0, \qquad \text{wenn } s < r.$$

Um sie zu beweisen, hat man nur die P-Produkte der v_i mit Hilfe von (D) als Schnitt der z_i zu schreiben, und dann die Rechenregeln für $\mathfrak{R}(G)$ anzuwenden.

10. Es seien $x = z_{i_1} \cdot z_{i_2} \cdot \cdots \cdot z_{i_r}$ und $y = z_{j_1} \cdot z_{j_2} \cdot \cdots \cdot z_{j_{l-r}}$ zwei Elemente der Basis $[z_i]$, und es sei $x \cdot y = \pm v_0$, d. h. die Elemente $z_{i_1}, \cdots, z_{i_r}, z_{j_1}, \cdots, z_{j_{l-r}}$ sind die Elemente $z_1, z_2, \cdots, z_l$, nur in anderer Reihenfolge. Auf Grund von (D) ist dann

$$x \circ y = v_{j_1} \circ v_{j_2} \circ \cdots \circ v_{j_{l-r}} \circ v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r};$$

und dieses Produkt ist gleich ± 1, weil die darin auftretenden v_i bis auf die Reihenfolge mit den Elementen $v_1, v_2, \cdots, v_l$ übereinstimmen. Führt man die elementare Bestimmung des Vorzeichens durch, so erhält man: Aus $x \cdot y = v_0$ folgt $x \circ y = 1$. Daraus folgt leicht allgemein: Sind X, Y zwei beliebige Elemente von $\mathfrak{B}(G)$ mit $d(X) + d(Y) = n$, dann ist die Schnittzahl von X und Y gleich ihrer "Aufspannzahl," d. h. in $X \cdot Y = a \cdot v_0$ und in $X \circ Y = b \cdot 1$ sind die Koeffizienten a und b gleich; diese Gleichheit ist ja anschaulich sehr naheliegend.

Eine letzte Bemerkung ist folgende:

Die Elemente $v_1, \cdots, v_r$ seien irgendwelche, linear unabhängige Elemente von $\mathfrak{B}(G)$; dann sind die Elemente

$$v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_s} \qquad (1 \leqq i_k \leqq r,\ i_1 < i_2 < \cdots < i_s)$$

linear unabhängige Elemente von $\mathfrak{B}(G)$. Zum Beweis hat man nur die $v_1, \cdots, v_r$ in eine Basis $v_1, \cdots, v_l$ von $\mathfrak{B}(G)$ aufzunehmen, und mit dieser Basis die Basis $\langle v_i \rangle$ von $\mathfrak{B}(G)$ zu bilden.

KAPITEL II

ERSTE ANWENDUNGEN DES AUFSPANNSATZES

Der Aufspannsatz soll jetzt mit den wichtigsten gruppentheoretischen Begriffen in Zusammenhang gebracht werden: mit dem der homomorphen Abbildung und hauptsächlich denen der Untergruppe und der Nebengruppenzerlegung.

1. Homomorphe Abbildung

1. G und G' seien zwei Gruppenmannigfaltigkeiten. Eine Abbildung h von G in G' heisst homomorphe Abbildung, wenn sie

a) eine homomorphe Abbildung im gewöhnlichen gruppentheoretischen Sinn ist, und

b) eine stetige Abbildung der Mannigfaltigkeit G in die Mannigfaltigkeit G' ist.

Die homomorphe Abbildung h bildet den Ring $\mathfrak{B}(G)$ homomorph in den Ring $\mathfrak{B}(G')$ ab; die additive Homomorphie ist klar; die Gleichung

$$h(x \circ y) = h(x) \circ h(y)$$

für zwei beliebige Homologieklassen x, y von G und ihre Bilder $h(x)$, $h(y)$ bei h beweist man etwa folgendermassen:

Man bilde das topologische Produkt $G \times G$ durch die Abbildung $f(p \times q) = h(pq)$ in G' ab. Ist F die Abbildung $F(p \times q) = pq$ von $G \times G$ in G, ist F' die

entsprechende Abbildung für G', und bezeichnet man mit g die Abbildung $g(p \times q) = h(p) \times h(q)$ von $G \times G$ in $G' \times G'$, dann kann man wegen $h(p \cdot q) = h(p) \cdot h(q)$ die Abbildung f sowohl als hF als auch als $F'g$ schreiben. Nun ist $F(x \times y) = x \circ y$, also $hF(x \times y) = h(x \circ y)$. Andererseits ist (nach Kap. I, §1, Nr. 8) $g(x \times y) = h(x) \times h(y)$, und folglich $F'g(x \times y) = h(x) \circ h(y)$. Damit ist die Behauptung bewiesen.

Ist speziell $h(x)$ homolog Null in G', so folgt, dass auch $h(x \circ y)$ homolog Null ist für beliebiges y; denn es ist $0' \circ z' = 0'$ ($0'$ die Nullklasse, z' eine beliebige Homologieklasse von G').

2. Der Kern einer homomorphen Abbildung eines Ringes bzw. einer Gruppe ist die Menge derjenigen Elemente, deren Bild die Null ist; der Kern ist ein Ideal bzw. eine (invariante) Untergruppe.

Wir betrachten die Gruppen $\mathfrak{B}(G)$ und $\mathfrak{B}(G')$ der minimalen Elemente von G und G'. Durch h wird $\mathfrak{B}(G)$ homomorph in $\mathfrak{B}(G')$ abgebildet (nach Kap. I, §1, Nr. 4); $\mathfrak{B}^0(G)$ sei der Kern dieser Abbildung. $\mathfrak{B}(G)$ werde als direkte Summe $\mathfrak{B}^0(G) + \mathfrak{B}^1(G)$ dargestellt; dann wird $\mathfrak{B}^1(G)$ isomorph abgebildet. $\{v_1, \cdots, v_{l^0}\}$ sei eine Basis von $\mathfrak{B}^0$, $\{v_{l^0+1}, \cdots, v_l\}$ eine von $\mathfrak{B}^1$; $\{v_1, \cdots, v_l\}$ ist dann eine Basis von $\mathfrak{B}(G)$. Die Elemente $v_1, \cdots, v_{l^0}$ werden auf Null abgebildet; die Elemente $h(v_{l^0+1}), \cdots, h(v_l)$ sind in G' linear unabhängig.

Wir betrachten die von v_0 und den Produkten

$$v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r} \qquad (i_1 < i_2 < \cdots < i_r)$$

gebildete Basis $\langle v_i \rangle$ von $\mathfrak{B}(G)$ (vgl. Satz I a)). Nach der letzten Bemerkung in Nr. 1 wird ein solches Produkt sicher auf 0 abgebildet, wenn auch nur ein Faktor v_{i_k} auf 0 abgebildet wird, wenn also wenigstens ein Index i_k aus der Reihe $1, 2, \cdots, l^0$ stammt. Andererseits werden diejenigen Elemente $v_{j_1} \circ v_{j_2} \circ \cdots \circ v_{j_s}$, in denen alle Indizes $> l^0$ sind, auf die Elemente

$$h(v_{j_1} \circ v_{j_2} \circ \cdots \circ v_{j_s}) = h(v_{j_1}) \circ h(v_{j_2}) \circ \cdots \circ h(v_{j_s})$$

abgebildet, und diese Elemente sind in G' linear unabhängig, weil die—nach Kap. I, §1, Nr. 4 minimalen—Elemente $h(v_j)$ $(j = l^0 + 1, \cdots, l)$ linear unabhängig sind (Kap. I, §3, Nr. 10).

Das bedeutet offenbar, dass der Charakter der Abbildung h durch folgenden Satz bestimmt ist:

Satz II. *Ist h eine homomorphe Abbildung von G in G', so ist der Kern der dadurch bewirkten homomorphen Abbildung von $\mathfrak{P}(G)$ in $\mathfrak{P}(G')$ das Ideal von $\mathfrak{P}(G)$, das erzeugt wird von denjenigen minimalen Elementen von $\mathfrak{R}(G)$, deren Bild bei h die Null ist.*

Der Restklassenring von $\mathfrak{P}(G)$ nach diesem Ideal—der ja durch h isomorph abgebildet wird—, ist isomorph dem von den Elementen $v_0, v_{l^0+1}, \cdots, v_l$ erzeugten Teilring von $\mathfrak{P}(G)$.

In jedem von Null verschiedenen Ideal von $\mathfrak{P}(G)$ ist das Element 1 (die Eins von $\mathfrak{R}(G)$) enthalten; denn wegen der Gleichung $v_1 \circ v_2 \circ \cdots \circ v_l = 1$ gibt es zu

jedem Element $x(\neq 0)$ ein y mit $x \circ y = 1$. Daher ist eine direkte Konsequenz des Satzes II das folgende.

KOROLLAR 1. *Das Bild $h(1)$ des Elementes 1 ist dann und nur dann ~ 0 in G', wenn durch h wenigstens ein minimales Element $(\neq 0)$ von G auf 0 abgebildet wird. Ist $h(1) \sim 0$, dann wird der ganze Ring $\mathfrak{P}(G)$ isomorph abgebildet.*

Daraus schliesst man nun weiter:

KOROLLAR 2. *Ist G durch f homomorph in G' abgebildet, und G' durch g homomorph in G'' abgebildet, ist weiter $f(G) \sim 0$ in G', und $g(G') \sim 0$ in G'', so ist $gf(G) \sim 0$ in G''* (G, G' bezeichnen hier die Einselemente von $\mathfrak{R}(G)$, $\mathfrak{R}(G')$).

Denn durch den Homomorphismus gf wird sogar der ganze Ring $\mathfrak{P}(G)$ isomorph abgebildet, weil f und g Isomorphismen sind.

3. In Anwendung dieser Begriffe soll jetzt die Abbildung $q(p) = p^2$ von G in sich, die also jedem Punkt p von G sein Quadrat p^2 zuordnet, untersucht werden.

Bezeichnet man mit f die Abbildung $f(p) = p \times p$ von G in $G \times G$ und mit F die schon betrachtete Abbildung $F(p \times p') = p \cdot p'$, so kann man die Abbildung $q(p)$ als zusammengesetzte Abbildung $Ff(p)$ schreiben. Das Produkt $G \times G$ ist selbst eine Gruppe, als direktes Produkt von G und G; und $f(p) = p \times p$ ist eine homomorphe Abbildung von G in $G \times G$.

Nach Nr. 2 wird man die f-Bilder der minimalen Elemente von G betrachten. Wir behaupten: Für jedes Element v von $\mathfrak{B}(G)$ ist

$$f(v) = (v \times v_0) + (v_0 \times v).$$

Zum Beweis beachten wir, dass $f(v)$ als minimales Element von $\mathfrak{R}(G \times G)$ von der Form $(v^1 \times v_0) + (v_0 \times v^2)$ mit $v^1, v^2 \in \mathfrak{B}(G)$ ist (Kap. I, §1, Nr. 9). Um zu zeigen, dass $v^1 = v^2 = v$ gilt, betrachten wir die Abbildungen $\pi_1(p \times p') = p$ und $\pi_2(p \times p') = p'$ von $G \times G$ in G; die Abbildungen $\pi_1 f$ und $\pi_2 f$ sind dann die Identität von G. Man erkennt sofort: Für jedes Element X der Form $(x \times v_0) + (v_0 \times y)$ ist $\pi_1(X) = x$ und $\pi_2(X) = y$; für jedes z von $\mathfrak{B}(G)$ ist $\pi_1 f(z) = \pi_2 f(z) = z$. Damit hat man die folgenden Gleichungen:

$$v = \pi_1 f(v) = \pi_1((v^1 \times v_0) + (v_0 \times v^2)) = v^1;$$

ebenso zeigt man: $v = v^2$, womit die obige Behauptung bewiesen ist.

Sei jetzt $\{v_1, v_2, \cdots, v_l\}$ eine Basis von $\mathfrak{B}(G)$. Nach dem eben Bewiesenen ist dann

$$f(v_i) = (v_i \times v_0) + (v_0 \times v_i),$$

also

$$q(v_i) = Ff(v_i) = F((v_i \times v_0) + (v_0 \times v_i)) = v_i + v_i = 2v_i.$$

Die Bilder $f(v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r})$ der Elemente der Basis $\langle v_i \rangle$ bestimmt man jetzt daraus, dass f homomorph ist. Man hat dabei zu beachten, dass für die Pontrjagin'sche Multiplikation in $G \times G$, wie man leicht sieht, gilt:

$$(x \times y) \circ (x' \times y') = (-1)^{d(y)d(x')}(x \circ x') \times (y \circ y').$$

Damit läuft die Rechnung so:

$$\begin{aligned} f(v_i \circ v_k) &= f(v_i) \circ f(v_k) \\ &= ((v_i \times v_0) + (v_0 \times v_i)) \circ ((v_k \times v_0) + (v_0 \times v_k)) \\ &= ((v_i \circ v_k) \times v_0) + (v_i \times v_k) - (v_k \times v_i) + (v_0 \times (v_i \circ v_k)); \end{aligned}$$

also

$$\begin{aligned} q(v_i \circ v_k) &= Ff(v_i \circ v_k) \\ &= F((v_i \circ v_k) \times v_0) + F(v_i \times v_k) - F(v_k \times v_i) + F(v_0 \times (v_i \circ v_k)) \\ &= (v_i \circ v_k) + (v_i \circ v_k) - (v_k \circ v_i) + (v_i \circ v_k); \end{aligned}$$

und wegen $v_i \circ v_k = -v_k \circ v_i$ erhält man endlich:

$$q(v_i \circ v_k) = 4 v_i \circ v_k .$$

Entsprechend verfährt man im allgemeinen Falle, bei der Berechnung von $q(v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r}) = Ff(v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r})$. Man geht aus von der Gleichung

$$f(v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r}) = f(v_{i_1}) \circ f(v_{i_2}) \circ \cdots \circ f(v_{i_r}),$$

setzt $f(v_{i_k}) = (v_{i_k} \times v_0) + (v_0 \times v_{i_k})$ ein, und multipliziert aus. Man erhält 2^r Summanden, von denen jeder das Element $\pm\, v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r}$ als Bild bei F hat—und man stellt fest, dass immer das positive Vorzeichen gilt. Damit hat man den Satz:

Der Homologietyp der Abbildung $q(p) = p^2$ von G in sich ist bestimmt durch die folgenden Gleichungen, in denen die q-Bilder der Elemente der Basis $\langle v_i \rangle$ von $\mathfrak{B}(G)$ angegeben sind:

$$q(v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r}) = 2^r v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r} \qquad (i_1 < i_2 < \cdots < i_r)$$

(*und natürlich* $q(v_0) = v_0$).

Darin ist, wegen der Gleichung $1 = v_1 \circ v_2 \circ \cdots \circ v_l$, enthalten:[5] Die Abbildung $q(p) = p^2$ von G in sich hat den Grad 2^l.

2. Untergruppen und Nebengruppenzerlegung

1. G sei wie immer eine Gruppenmannigfaltigkeit. Eine Teilmenge U heisst eine Untergruppe, wenn sie

a) eine Untergruppe von G im gewöhnlichen gruppentheoretischen Sinn ist, und

b) eine abgeschlossene Teilmenge von G ist.

Dazu bemerken wir folgendes: Es ist bekannt, dass jede geschlossene Gruppenmannigfaltigkeit eine Lie'sche Gruppe, also jedenfalls einige Male differenzierbar ist. Ueber das Verhalten einer Untergruppe U ist bekannt [17]: Man kann in einer hinreichend kleinen, dem Euklidischen R_n ($n = d(G)$) homöomorphen,

[5] Vgl.: H. Hopf: *Über den Rang geschlossener Liescher Gruppen* [Comm. Math. Helv 13. (1940/41)], Satz I für $k = 2$.

Umgebung V des Einheitspunktes e der Gruppe G ein solches Koordinatensystem einführen, dass diejenigen Punkte von U, die zu V gehören, die also den Durchschnitt $U \cdot V$ bilden, eine gewisse Ebene durch e erfüllen. U ist also im kleinen euklidisch. Beachtet man noch, dass U als abgeschlossene Teilmenge von G kompakt ist, so sieht man leicht: Ist die Untergruppe U zusammenhängend, dann ist sie eine differenzierbare und differenzierbar in G eingelagerte, geschlossene, Gruppenmannigfaltigkeit; ist U nicht zusammenhängend, dann ist die Komponente U', die den Punkt e enthält, selbst eine zusammenhängende Untergruppe von G, also eine Gruppenmannigfaltigkeit in G, und U besteht aus endlich viel Komponenten, die Nebengruppen von U' sind.

Deshalb bedeutet es kaum eine Einschränkung, wenn wir im folgenden die Forderung b) ersetzen durch die Forderung

b′) U ist eine (differenzierbar) in G eingelagerte geschlossene Mannigfaltigkeit (also zusammenhängend).

Da U eine Gruppenmannigfaltigkeit ist, haben für sie die Definitionen und Sätze des Kap. I Gültigkeit.

2. G sei eine Gruppe, U eine Untergruppe. Die Gruppenmannigfaltigkeit U ist dadurch, dass sie eine Untergruppe von G ist, auf natürliche Weise homomorph, sogar isomorph, in G abgebildet; man hat dazu die Punkte von U, die ja auch Punkte von G sind, eben als Punkte von G aufzufassen. Das Bild einer Homologieklasse u von U bei dieser Abbildung ist u, als Homologieklasse von G betrachtet; genauer: das Bild von u ist die Homologieklasse von G, in der die Zyklen aus u, die ja auch Zyklen in G sind, liegen. Die Homomorphie besteht darin, dass für die Homologieelemente von U die Pontrjagin'sche Multiplikation in U übereinstimmt mit der in G.

Man kann also die Sätze des §1 anwenden. Wir stellen die Frage: Welche Elemente von $\mathfrak{B}(U)$ sind homolog 0 in G? Nach §1 hängt das davon ab, welche *minimalen* Elemente von $\mathfrak{R}(U)$ in G homolog 0 sind. Aus dem Satz II folgt der

Satz III. *Die Gesamtheit der Homologieelemente der Untergruppe U, die in der Gruppe G homolog 0 sind, ist das Ideal von $\mathfrak{P}(U)$, das erzeugt wird von denjenigen minimalen Elementen von $\mathfrak{R}(U)$, die in G homolog 0 sind.*

Das sind also die Elemente $\sum w_h \circ v_h$, wo die w_h beliebige Elemente von $\mathfrak{P}(U)$ sind, und die v_h Elemente von $\mathfrak{B}(U)$ bedeuten, die in G homolog 0 sind.

Daraus entnimmt man sofort das

Korollar 1. *Eine Untergruppe U ist dann und nur dann ~ 0 in der Gruppe G, wenn wenigstens ein von Null verschiedenes minimales Element aus $\mathfrak{R}(U)$ in G homolog 0 ist.*

Weiter gilt:

Korollar 2. *Ist U nicht homolog 0 in G, dann ist kein von Null verschiedenes Element von $\mathfrak{B}(U)$ homolog 0 in G.*

Daraus wieder folgt sofort:

Korollar 3. *Ist G'' eine Untergruppe von G', die nicht homolog 0 ist in G', und ist G' eine Untergruppe von G, die nicht homolog 0 ist in G, dann ist G'' nicht homolog 0 in G.*

Dann ist sogar kein Element von $\mathfrak{B}(G'')$ homolog 0 in G.

Diesen Satz verallgemeinert man zu folgendem:

Ist $G_1 \supset G_2 \supset G_3 \supset \cdots \supset G_r$ *eine absteigende Kette von Untergruppen, und ist* $G_{i+1} \nsim 0$ *in* G_i *(für* $i = 1, 2, \cdots, r-1$*), dann ist* $G_r \nsim 0$ *in* G_1.

Man kann die Ueberlegungen, die zum Satz III führen, etwas anders fassen:

$\mathfrak{V}^0(U)$ sei die Gruppe derjenigen Elemente von $\mathfrak{B}(U)$, die ~ 0 in G sind. Man stelle $\mathfrak{B}(U)$ als direkte Summe $\mathfrak{V}^0(U) + \mathfrak{V}^1(U)$ dar. Die Elemente $v_1, v_2, \cdots, v_r$ einer Basis von $\mathfrak{V}^1(U)$ sind dann in G linear unabhängig; man kann sie sowohl in eine Basis von $\mathfrak{B}(U)$ als auch in eine Basis von $\mathfrak{B}(G)$ aufnehmen. Die Produkte $v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_s}$ $(i_1 < i_2 < \cdots < i_r,\ 1 \leqq i_k \leqq r)$, die man sowohl als Elemente von $\mathfrak{P}(U)$ als auch als Elemente von $\mathfrak{P}(G)$ auffassen kann, sind nach dem Aufspannsatz sowohl in $\mathfrak{P}(U)$ als auch in $\mathfrak{P}(G)$ linear unabhängig; sie bilden eine Basis der Gruppe derjenigen Homologieklassen von G, in denen Homologieklassen von U liegen; die Elemente von $\mathfrak{B}(U)$, die ~ 0 in G sind, sind die Elemente $\sum w_h \circ v_h$ mit $w_h \subset \mathfrak{B}(U)$, $v_h \subset \mathfrak{V}^0(U)$.

Gleichbedeutend damit, dass $U \nsim 0$ in G ist, ist: $\mathfrak{V}^0(U)$ enthält nur die Null.

3. Ist G eine Gruppe, U eine Untergruppe, so zerfällt G in die (etwa linksseitigen) Nebengruppen, was man durch die Gleichung

$$G = U + a_1 \cdot U + a_2 \cdot U + \cdots$$

ausdrückt, in der die a_i gewisse Elemente von G sind. (Wir gebrauchen die übliche Schreibweise auch hier, wo die Anzahl der Nebengruppen kontinuierlich ist.)

Die Wichtigkeit dieser Nebengruppenzerlegung für topologische Gruppen G besteht darin, dass man die Nebengruppen als Punkte eines topologischen Raumes, des Nebengruppenraumes, auffassen kann, dessen Topologie eng mit der von G verknüpft ist. Dieser Raum, der mit G/U oder mit W ("Wirkungsraum," vgl. Nr. 5) bezeichnet wird, ist also folgendermassen definiert:

Die Punkte von W sind die Nebengruppen aU von U; eine Umgebung eines Punktes aU besteht aus den Nebengruppen $a' \cdot U$, wo der Punkt a' eine Umgebung von a in G durchläuft.

Auf Grund der Bemerkungen in Nr. 1 ist leicht zu sehen, dass auch der Nebengruppenraum eine differenzierbare Mannigfaltigkeit ist. Man führt dazu in einer (hinreichend kleinen) Umgebung des beliebigen Punktes a von G ein Koordinatensystem so ein, dass die Nebengruppe aU durch eine r-dimensionale Ebene $(r = d(U))$ durch a dargestellt wird; man stellt leicht fest, dass die Nebengruppen, die eine Umgebung von aU in W bilden, eineindeutig den Punkten der zu aU senkrechten, $(n - r)$-dimensionalen Ebene durch a entsprechen $(n = d(G)$. Damit hat man in der Umgebung des Punktes aU von W ein Euklidisches Koordinatensystem eingeführt. Die Differenzierbarkeit zeigt man ebenso.

Wir könne und wollen also im folgenden annehmen, dass der Nebengruppen-

raum W eine differenzierbare, also triangulierbare [18], geschlossene Mannigfaltigkeit ist.

4. Die Zerlegung von G in die Nebengruppen aU ist eine Faserung mit der Faser U; dabei versteht man unter der Faserung einer Mannigfaltigkeit M mit der Faser (-Mannigfaltigkeit) F folgendes [10]: M ist zerlegt in Mannigfaltigkeiten, die alle einer festen Mannigfaltigkeit F homöomorph sind, und die "Fasern" von M heissen; durch jeden Punkt von M geht also eine und nur eine Faser; und eine Umgebung einer beliebigen Faser ist fasertreu homöomorph dem topologischen Produkt $E \times F$, wo E ein Element (Vollkugel) der Dimension $d(E) = d(M) - d(F)$ ist.

"Fasertreu homöomorph" heisst dabei: jede Faser der Umgebung ist topologisch auf eine der Mannigfaltigkeiten $p \times F$ abgebildet (p ein Punkt von E)—verschiedene Fasern natürlich auf verschiedene $p \times F$.

Fasst man die Fasern als Punkte eines Raumes, mit natürlichem Umgebungsbegriff, auf (bildet man den sog. Zerlegungsraum der Zerlegung von M in die Fasern), so erhält man den Faser- oder Basisraum B, der offenbar eine Mannigfaltigkeit ist; der Faserraum wird auch, in einer naheliegenden Schreibweise, mit M/F bezeichnet.

Wir sprechen im folgenden von Faserungen auch bei Teilmengen einer gefaserten Mannigfaltigkeit, die ganz aus Fasern bestehen, d. h. zu einer beliebigen Faser entweder fremd sind oder sie ganz enthalten; und ist K ein beliebiger Komplex, so nennen wir das topologische Produkt $K \times F$ gefasert—die Fasern sind natürlich die Mengen $p \times F$, wo p ein beliebiger Punkt von K ist. Es ist dann klar, was man unter einer fasertreuen (topologischen) Abbildung zu verstehen hat.

Ordnet man jedem Punkt einer gefaserten Mannigfaltigkeit M die Faser zu, auf der er liegt, so erhält man eine stetige Abbildung von M auf den Faserraum B, die Projektion P. Das Urbild $P^{-1}(T)$ einer Teilmenge T von B ist die (gefaserte) Menge der Punkte der "über T stehenden" Fasern. Für jede hinreichend kleine Umgebung V eines beliebigen Punktes von B ist $P^{-1}(V)$ fasertreu homöomorph mit $V \times F$; M ist im kleinen ein topologisches Produkt. Ueber dieses lokale Zerfallen in ein topologisches Produkt gilt der folgende Satz von Feldbau [19]:

Ist E ein Element (Vollkugel) im Faserraum B, dann ist die Urbildmenge $P^{-1}(E)$ dem Produkt $E \times F$ fasertreu homöomorph.

Nach Nr. 3 ist leicht zu sehen, dass die Zerlegung einer Gruppe G in die Nebengruppen aU einer Untergruppe U eine Faserung von G mit der Faser U ist. Man erkennt nämlich sofort: Durchläuft p die in Nr. 3 genannte, zu aU senkrechte Ebene, die E heissen möge, in der Umgebung von a, und durchläuft q die Untergruppe U, dann ist $f(p \times q) = p \cdot q$ eine fasertreue topologische Abbildung von $E \times U$ auf die Umgebung von U (pq ist das Produkt von p und q in G).

Der Faserraum dieser Faserung ist der Nebengruppen- oder Wirkungsraum W. Die Projektion P besteht darin, dass man dem Punkt p von G die Nebengruppe pU (als Punkt von W) zuordnet.

5. Eine Mannigfaltigkeit W heisst Wirkungsraum der Gruppenmannigfaltigkeit G ("espace homogène") [9], wenn folgendes erfüllt ist:

Jedem Punkt p von G ist eine topologische Transformation T_p von W auf sich zugeordnet; die Transformationen T_p hängen stetig von dem Punkt p ab; es gilt: $T_p \cdot T_q = T_{pq}$; und die aus den T_p bestehende Transformationsgruppe von W ist transitiv, d. h. jeder Punkt von W (es genügt: ein bestimmter Punkt von W) kann (durch geeignete Transformationen T_p) in jeden Punkt von W übergeführt werden.

a sei ein Punkt von W; diejenigen Punkte p von G, für die $T_p(a) = a$ gilt, d. h. deren T_p den Punkt a als Fixpunkt hat, bilden eine Untergruppe von G, die "Isotropiegruppe" von W. (Setzt man $W = R_n$ und nimmt als G die Gruppe der Bewegungen des R_n , so ist die Isotropiegruppe die Gruppe der Rotationen um einen festen Punkt.) Die zu verschiedenen Punkten a von W gehörigen Isotropiegruppen sind in G konjugiert, sodass man von *der* Isotropiegruppe U sprechen kann.

Die Punkte p von G, deren T_p den Punkt a in den gleichen Punkt b transformieren, bilden eine linke Nebengruppe von U. Diese eineindeutige Zuordnung der Punkte von W und der linken Nebengruppen von U ist eine Homöomorphie zwischen W und dem Nebengruppenraum G/U.

Die Transformationen T_p können also auch als Transformationen des Nebengruppenraumes G/U aufgefasst werden; sie haben da eine einfache Bedeutung: Die Transformation T_p besteht nämlich darin, dass man jeder Nebengruppe qU die Nebengruppe pqU zuordnet.

Ein Wirkungsraum W ist also identisch mit dem zur Isotropiegruppe U von W gehörigen Nebengruppenraum G/U (bei der Definition des Nebengruppenraums muss man dann mehrkomponentige Untergruppen (vgl. Nr. 1) U zulassen).

Diese Definition des Wirkungsraumes ist etwas allgemeiner als die übliche: wir lassen zu, dass T_p für gewisse p die Identität von W ist (p gehört dann zu U).

Ist U zusammenhängend, dann ist W orientierbar (aber nicht nur dann): Man stelle (nach Nr. 4) die Umgebung einer Nebengruppe pU als topologisches Produkt $E \times U$ dar. Die Orientierung von G und die von U bestimmen dann eindeutig eine Orientierung von E, und diese überträgt man mittels der Projektion P in den Wirkungsraum W. (Entsprechend für Faserungen M/F.)

6. Das Problem, das sich jetzt erhebt, ist: Zusammenhänge zwischen G, U, W, insbesondere zwischen ihren Homologieeigenschaften, zu finden. Es handelt sich z. B. um die Aufgabe, aus den Homologieeigenschaften von G und U die von W zu bestimmen, wobei man natürlich die Lage von U in G berücksichtigen muss. Zur Lösung dieser Aufgabe werden in Kap. IV einige Beiträge geliefert. Noch naheliegender ist die Fragestellung: was kann man bei gegebenem G und W über U aussagen?, also die Frage nach der Isotropiegruppe U eines vorgelegten Wirkungsraumes W. Diese Frage wird im Kap. III für den speziellen Fall $W = S_n$ gelöst; als Anwendung werden die Homologieringe der Gruppen A_n , B_n , C_n , D_n bestimmt.

Kapitel III

Die Sphären als Wirkungsräume

In diesem Kapitel wird der Spezialfall untersucht, dass der zu der Gruppe G und der Untergruppe U gehörige Wirkungsraum W der m-dimensionalen Sphäre S_m homöomorph ist. Dieser Fall tritt auf bei den Gruppen der bekannten vier Klassen A_n, B_n, C_n, D_n; die Homologieringe dieser Gruppen lassen sich damit vollständig bestimmen.

1. $G/U = S_m$

1. G sei eine Gruppe; der Rang von G (Kap. I, §2, Nr. 3) sei l; U sei eine Untergruppe von G. Wir beweisen den folgenden Satz:

Satz IV. *Der zu der Untergruppe U der Gruppe G gehörige Wirkungsraum $G/U = W$ sei der m-dimensionalen Sphäre S_m homöomorph; dann gilt:*

a) *ist m ungerade, so ist $\mathfrak{R}(G)$ isomorph mit $\mathfrak{R}(U \times S_m)$, und U ist ~ 0 in G,*

b) *ist m gerade, so ist $\mathfrak{R}(U)$ isomorph mit $\mathfrak{R}(\Pi \times S_{m-1})$, wo Π ein topologisches Produkt von $l-1$ Sphären ungerader Dimensionen ist, und $\mathfrak{R}(G)$ ist isomorph mit $\mathfrak{R}(\Pi \times S_{2m-1})$; und U ist ~ 0 in G.*

Bemerkung: für $l = 1$ bedeutet b): $\mathfrak{R}(U) \approx \mathfrak{R}(S_{m-1})$, und $\mathfrak{R}(G) \approx \mathfrak{R}(S_{2m-1})$. (Dann ist übrigens, was aber für das folgende unwesentlich ist, notwendigerweise $m = 2$.)

Für die Ränge $l(G)$, $l(U)$ und die charakteristischen Zahlen $l_i(G)$, $l_i(U)$ (Kap. I, §2, Nr. 3) bedeutet dieser Satz das Bestehen der folgenden Gleichungen:

a) m ungerade:

$$\begin{aligned} l(G) &= l(U) + 1, \\ l_m(G) &= l_m(U) + 1, \\ l_i(G) &= l_i(U) \qquad \text{für } i \neq m, \end{aligned}$$

b) m gerade:

$$\begin{aligned} l(G) &= l(U) \\ l_{m-1}(G) &= l_{m-1}(U) - 1, \\ l_{2m-1}(G) &= l_{2m-1}(U) + 1, \\ l_i(G) &= l_i(U) \qquad \text{für } i \neq m-1, \neq 2m-1. \end{aligned}$$

Der Beweis beruht auf der Betrachtung der Faserung von G in die Nebengruppen aU von U.

2. Wir erinnern zunächst an einige Begriffe:

Ein endliches, euklidisches Polyeder ist die Menge der Punkte, die zu einem endlichen, euklidischen, simplizialen (allgemeiner: Zellen-) Komplex gehören. Ein endliches, krummes Polyeder (kurz: Polyeder) P in einer Mannigfaltigkeit

M ist das topologische Bild in M eines endlichen, euklidischen Polyeders [20]. Ein stetiger Zyklus ist das stetige Bild eines, in einem euklidischen R_n liegenden, simplizialen "Parameter"-Zyklus. Ein stetiger Zyklus z in einem Polyeder P bestimmt eine Homologieklasse von P, die man eben durch z repräsentieren kann [21]. Im folgenden verstehen wir unter Zyklen immer stetige Zyklen.

Wir betrachten eine beliebige Gruppe G. Im Kap. I, §3, Nr. 1 ist mit Hilfe der Abbildung $F(p \times q) = pq$ von $G \times G$ in G das Pontrjagin'sche Produkt $x \circ y$ zweier Homologieklassen x und y von G definiert worden. Diese Definition lässt sich unmittelbar erweitern zu der des Pontrjagin'schen Produktes $z_1 \circ z_2$ zweier stetiger Zyklen z_1, z_2 in G: man bilde den stetigen Zyklus $z_1 \times z_2$ in $G \times G$, und definiere $z_1 \circ z_2$ durch

$$z_1 \circ z_2 = F(z_1 \times z_2);$$

$z_1 \circ z_2$ ist ebenfalls ein stetiger Zyklus in G.

3. Jetzt sei gemäss der Voraussetzung von Satz IV der zur Untergruppe U der Gruppe G gehörige Nebengruppenraum G/U der S_m homöomorph. P bezeichnet die Projektion von G auf $G/U = S_m$.

Wir zerlegen die S_m durch eine $(m-1)$-dimensionale Aequatorsphäre S_{m-1} in zwei Elemente (Vollkugeln) E^1 und E^2, die "nördliche und südliche Halbkugel." Es ist also $E^1 + E^2 = S_m$ und $E^1 \cdot E^2 = S_{m-1}$. Das Urbild $P^{-1}(E^1)$ von E^1 heisse G', das Urbild $P^{-1}(E^2)$ von E^2 heisse G''. Es ist $G = G' + G''$; d. h., G wird von G' und G'' überdeckt. Der Durchschnitt $G' \cdot G''$, der H heisse, ist das Urbild $P^{-1}(S_{m-1})$ von S_{m-1}.

Wir wenden jetzt den Satz von Feldbau an (Kap. II, §2, Nr. 4). Weil E^1 ein Element ist, ist danach die Menge G', die ja ganz aus Nebengruppen von U besteht, dem topologischen Produkt $E^1 \times U$ fasertreu homöomorph. Dabei ist nur benutzt, dass eine Faserung von G in die Nebengruppen von U vorliegt. Indem man nun die Gruppeneigenschaft stärker ausnützt, kann man die Darstellung von G' als Produkt $E^1 \times U$ in einer speziellen Weise wählen. Man betrachte dazu für einen beliebig, aber fest gewählten Punkt q von U das in $E^1 \times U$ liegende Element $E^1 \times q$. Ihm entspricht infolge der Homöomorphie von $E^1 \times U$ und G' ein Element E' in G'. Aus der Definition folgt unmittelbar, dass E' mit jeder zu G' gehörenden Faser von G (Nebengruppe von U) genau einen, und mit jeder nicht zu G' gehörenden Faser keinen Punkt gemeinsam hat und dass E' durch P topologisch auf E^1 abgebildet wird. Man betrachte nun das in $G \times G$ liegende Polyeder $E' \times U$. Bei der Abbildung F wird dieses Polyeder topologisch-fasertreu auf G' abgebildet. Denn ist p ein beliebiger Punkt von E', so wird die Teilmenge $p \times U$ von $E' \times U$ topologisch auf die, in G' enthaltene, Nebengruppe pU abgebildet (das ist geradezu die Definition der Nebengruppe); sind p_1, p_2 zwei verschiedene Punkte von E', so sind, wie eben bemerkt, die Nebengruppen $p_1 \cdot U$ und $p_2 \cdot U$, auf die die Mengen $p_1 \times U$ und $p_2 \times U$ abgebildet werden, verschieden; und man erhält so auch alle zu G' gehörenden Nebengruppen. Die ursprüngsliche Homöomorphie zwischen

$E^1 \times U$ und G' ist also jetzt ersetzt durch die durch F vermittelte Homöomorphie zwischen $E' \times U$ und G'. Entsprechend lässt sich G'' darstellen als topologisches F-Bild eines in $G \times G$ liegenden Produktes $E'' \times U$, wo E'' ein Element in G ist, das durch P topologisch auf E^2 abgebildet wird. G' und G'' sind also krumme Polyeder in G; sie sind sogar berandete Mannigfaltigkeiten.

Die Randsphäre von E' heisse S', die von E'' heisse S''. Beide Sphären werden durch P topologisch auf S_{m-1} abgebildet. Dann sieht man: Bei der topologisch-fasertreuen Abbildung F von $E' \times U$ auf G' erscheint als Bild des Randes $S' \times U$ von $E' \times U$ die Menge $H = P^{-1}(S_{m-1}) = G' \cdot G''$. Denn $S' \times U$ wird bei F topologisch auf die Menge der Punkte derjenigen Nebengruppen von U abgebildet, die einen Punkt mit S' gemeinsam haben, deren Projektion, d. h. deren Bild bei P, also zu S_{m-1} gehört. Als topologisches Bild der Mannigfaltigkeit $S' \times U$ ist H eine Mannigfaltigkeit in G.

Ebenso stellt man fest, dass H topologisch-fasertreues F-Bild des Randes $S'' \times U$ von $E'' \times U$ ist.

4. Den orientierten Grundzyklus der Sphäre S' bezeichnen wir auch mit S'; p' sei ein einfach gezählter Punkt von S'; die Zyklen $u_1, \cdots, u_r$ mögen eine Basis von $\mathfrak{B}(U)$ bilden. Dann bilden die Zyklen $p' \times u_i$ und $S' \times u_i$ $(i = 1, \cdots, r)$ eine Basis von $\mathfrak{B}(S' \times U)$; für das topologische Bild $H = F(S' \times U)$ bedeutet das: die (nach Nr. 2 gebildeten) Zyklen

$$p' \circ u_i, \qquad S' \circ u_i \qquad (i = 1, \cdots, r)$$

bilden eine Basis von $\mathfrak{B}(H)$; diese Basis heisse $\mathfrak{B}_1$. Ebenso folgt aus der Tatsache, dass H topologisches F-Bild von $S'' \times U$ ist, dass die Zyklen

$$p'' \circ u_i, \qquad S'' \circ u_i \qquad (i = 1, \cdots, r)$$

eine Basis von $\mathfrak{B}(H)$ bilden; p'', S'' sind analog wie p', S' definiert.

Man sieht nun leicht ein, dass man die beiden Punkte p' und p'' und ebenso den einfach gezählten Punkt p_0 von $\mathfrak{B}(U)$ durch den Einheitspunkt e von G repräsentieren kann. (Man kann zunächst S_{m-1} so wählen, dass U in $P^{-1}(S_{m-1}) = H$ liegt, und kann dann die Elemente E' und E'' so bestimmen, dass sie den Punkt e in ihrem Rand enthalten.)

Man betrachte die Untergruppe von $\mathfrak{B}(H)$, die von den Elementen $p' \circ u$ mit beliebigem u aus $\mathfrak{B}(U)$ gebildet wird; sie ist mit $\mathfrak{B}(U)$ isomorph (entsprechend der Tatsache, dass die Gruppe der Elemente $p' \times u$ von $\mathfrak{B}(S' \times U)$ mit $\mathfrak{B}(U)$ isomorph ist). Wegen der speziellen Wahl von p' ist

$$p' \circ u = e \circ u = u;$$

das bedeutet: man kann die Elemente $p' \circ u$ dieser Untergruppe mit den Elementen u von $\mathfrak{B}(U)$ identifizieren, man kann also $\mathfrak{B}(U)$ als Untergruppe von $\mathfrak{B}(H)$ auffassen. Die Elemente $S' \circ u$ (und ebenso die Elemente $S'' \circ u$) gehören nicht zu dieser Untergruppe. Wir bemerken schliesslich noch, dass wir, wegen $p_0 = e$, das Element $S' \circ p_0$ von $\mathfrak{B}(H)$ mit der Sphäre S' identifizieren können; ebenso identifizieren wir $S'' \circ p_0$ mit S''.

5. Wir stellen jetzt die Frage: Welche Elemente von $\mathfrak{B}(U)$ sind homolog 0 in G? Um sie zu beantworten, verallgemeinern wir sie zu der Frage: Welche Elemente von $\mathfrak{B}(H)$ sind homolog 0 in G?

Wir gehen aus von der Tatsache, dass G überdeckt ist mit den Polyedern G' und G'', deren Durchschnitt eben das Polyeder H ist. Daher lässt sich ein elementarer Satz aus dem Kreis der "Additionssätze der kombinatorischen Topologie" [22] anwenden, der besagt:

Die Homologieklasse z von H ist dann und nur dann homolog 0 in G, wenn sie sich in H darstellen lässt als Summe $z_1 + z_2$ zweier Homologieklassen z_1, z_2 von H, wobei z_1 in G' und z_2 in G'' homolog 0 ist.

(Im rein kombinatorischen Fall verläuft der Beweis dieses Satzes folgendermassen: der in H liegende Zyklus z sei ~ 0 in G, also $z = \dot{C}$. (Mit $\dot{C}$ bezeichnen wir den Rand des Komplexes C.) Unter C_1 verstehe man den Teilkomplex von C, dessen Simplexe in G' liegen, unter C_2 den von den übrigen Simplexen von C gebildeten Teilkomplex von C; es ist also $C = C_1 + C_2$. Dann ist $z = \dot{C} = \dot{C}_1 + \dot{C}_2$, und $\dot{C}_1$, bzw. $\dot{C}_2$, ist ein Zyklus in H, der ~ 0 in G', bzw. G'', ist. Im vorliegenden Fall krummer Polyeder muss man die vorkommenden Zyklen und Komplexe als stetige Zyklen und stetige Komplexe annehmen und noch einige einfache Approximationen vornehmen.)

6. Die Zyklen von H, die in G' homolog 0 sind, sind nun genau die Zyklen $S' \circ u$, wo u ein beliebiges Element von $\mathfrak{B}(U)$ ist, entsprechend der Tatsache, dass in $E' \times U$ genau die Homologieklassen $S' \times u$ des Randes $S' \times U$ von $E' \times U$ homolog 0 sind. Analog sind genau die Elemente $S'' \circ u$ von $\mathfrak{B}(H)$ in $G'' = F(E'' \times U)$ homolog 0. Die Elemente R von $\mathfrak{B}(H)$, die in G homolog 0 sind, sind also nach dem genannten Additionssatz die Elemente

$$R = S' \circ u_1 + S'' \circ u_2 \tag{1}$$

mit beliebigem u_1 und u_2 aus $\mathfrak{B}(U)$.

Um eine Uebersicht über sie zu bekommen, schreiben wir sie in der Basis $\mathfrak{B}_1$ von $\mathfrak{B}(H)$. Dazu stellen wir zunächst einmal S'' in dieser Basis dar:

$$S'' \sim S' \circ u' + p' \circ u \qquad \text{in } H.$$

Aus Dimensionsgründen ist u' gleich cp_0 mit einem gewissen Koeffizienten c. Wegen der in Nr. 4 getroffenen speziellen Wahl von p_0 und p' hat man also:

$$S'' \sim cS' + u \qquad \text{in } H.$$

S' und S'' sind als Sphären minimale Elemente von $\mathfrak{R}(H)$ (vgl. Kap. I, §1, Nr. 4); da die minimalen Elemente eine Gruppe bilden, ist auch u ein minimales Element von $\mathfrak{R}(H)$. Aus dem Zusammenhang zwischen den minimalen Elementen eines topologischen Produktes mit den minimalen Elementen seiner Faktoren (Kap. I, §1, Nr. 9) folgt, dass u sogar ein minimales Element von $\mathfrak{R}(U)$ ist, das wir mit $\bar{v}$ bezeichnen wollen. Es ist also schliesslich:

(2) $$S'' \sim cS' + \bar{v} \qquad \text{in } H.$$

Dabei ist $\bar{v}$ eventuell das Nullelement.

(2) in (1) eingesetzt, ergibt die folgende Gleichung:

$$R \sim S' \circ (u_1 + cu_2) + \bar{v} \circ u_2 \qquad \text{in } H.$$

(Dabei ist benützt, dass für einen Zyklus z in H und einen Zyklus u in U das Produkt $z \circ u$ in H liegt, und dass für zwei Zyklen z_1, z_2 in H und einen Zyklus u in U gilt:

$$(z_1 + z_2) \circ u \sim z_1 \circ u + z_2 \circ u \qquad \text{in } H.$$

Um das zu beweisen, beachte man, dass H aus linken Nebengruppen von U besteht, und dass folglich die Teilmenge $H \times U$ von $G \times G$ durch F in H abgebildet wird. In $H \times U$ ist nun $(z_1 + z_2) \times u$ homolog $(z_1 \times u) + (z_2 \times u)$; übt man darauf F aus, so erhält man die behauptete Homologie.)

Damit hat man schliesslich:

Die Gesamtheit der Elemente von $\mathfrak{B}(H)$, die in G homolog 0 sind, wird gebildet von den Elementen

(3) $$S' \circ u + \bar{v} \circ u_2 ,$$

wo u, u_2 beliebige Elemente von $\mathfrak{B}(U)$ sind.

Unter den Elementen (3) sind nur die mit $u = 0$ in der Gruppe $\mathfrak{B}(U)$ enthalten; also folgt:

Die Gesamtheit der Elemente von $\mathfrak{B}(U)$, die in G homolog 0 sind, wird gebildet von den Elementen

$$\bar{v} \circ u$$

mit beliebigem u aus $\mathfrak{B}(U)$, sie ist also das von $\bar{v} = S'' - cS'$ erzeugte Ideal von $\mathfrak{B}(U)$.

7. Nach Satz III (Kap. II, §2, Nr. 2) ist daher $\bar{v}$ das einzige minimale Element von $\mathfrak{R}(U)$, das homolog 0 in G ist; genauer: die Gruppe $\mathfrak{B}^0(U)$ derjenigen minimalen Elemente von $\mathfrak{R}(U)$, die in G homolog 0 sind, besteht aus den rationalen Vielfachen von $\bar{v}$.

Aus der Definition von $\bar{v}$ folgt: $d(\bar{v}) = d(S') = d(S'') = m - 1$. Wir unterscheiden jetzt die Fälle $\bar{v} = 0$ und $\bar{v} \neq 0$. Der Kürze halber wird der Rang von $\mathfrak{B}(G)$ mit l und der Rang von $\mathfrak{B}(U)$ mit l' bezeichnet.

a) Es sei $\bar{v} = 0$. Dann ist also (ausser 0) kein minimales Element von $\mathfrak{R}(U) \sim 0$ in G; also ist nach Kap. II, §2, Nr. 2, Korollar 1 die Untergruppe $U \nsim 0$ in G, und nach Korollar 2 kann man $\mathfrak{B}(U)$ als Untergruppe von $\mathfrak{B}(G)$ auffassen. Eine Basis $\{v_1, v_2, \cdots, v_{l'}\}$ von $\mathfrak{B}(U)$ kann man aufnehmen in eine Basis $\{v_1, v_2, \cdots, v_{l'}, v_{l'+1}, \cdots, v_l\}$ von $\mathfrak{B}(G)$. Wegen $\sum_1^{l'} d(v_i) = d(U)$, $\sum_1^{l} d(v_i) = d(G)$ und $d(G) = d(U) + m$ gilt dabei:

$$d(v_{l'+1}) + d(v_{l'+2}) + \cdots + d(v_l) = m.$$

Nun ist aber jede Homologieklasse von G, deren Dimension kleiner als m ist, in G einer Homologieklasse von U homolog. Denn aus der Tatsache, dass E^i ein m-dimensionales Element ist ($i = 1, 2$), schliesst man ohne weiteres, dass man jeden höchstens $(m-1)$-dimensionalen Zyklus von G auf den gemeinsamen Rand H von G' und G'' deformieren kann. Aber die Homologieelemente von $\mathfrak{B}(H)$, die ~ 0 in G sind, sind in G den Elementen der Untergruppe $\mathfrak{B}(U)$ von $\mathfrak{B}(H)$ homolog (weil die Elemente $S' \circ u$ von $\mathfrak{B}(H)$ in G homolog 0 sind).

Die Dimensionen der nicht zu der Untergruppe $\mathfrak{B}(U)$ von $\mathfrak{B}(G)$ gehörigen Elemente $v_{l'+1}, v_{l'+2}, \cdots, v_l$ müssen also $\geqq m$ sein; da die Summe der Dimensionen, wie oben festgestellt, gleich m ist, gibt es nur ein einziges solches Element, und es hat die Dimension m.

Das bedeutet: Man erhält eine Basis von $\mathfrak{B}(G)$, indem man zu einer Basis $\{v_1, v_2, \cdots, v_{l'}\}$ von $\mathfrak{B}(U)$ ein gewisses Element v_l mit $d(v_l) = m$ hinzufügt.

Im Falle eines ungeraden m ist nun das Element $\bar{v}$ ein minimales Element der geraden Dimension $m-1$, und folglich gleich Null (Kap. I, §2, Nr. 3). Die eben angestellten Ueberlegungen sind also anwendbar auf diesen Fall, und enthalten offenbar einen Beweis der Behauptung a) aus Nr. 1.

Da aus $\bar{v} = 0$ die Existenz eines minimalen Elementes $v_l \neq 0$ mit $d(v_l) = m$ folgt, und andererseits ein minimales Element gerader Dimension immer Null ist, so muss im Fall eines geraden m das Element $\bar{v}$ von Null verschieden sein. Damit kommen wir zum Fall b).

b) Es sei $\bar{v} \neq 0$. Nimmt man $\bar{v}$ in eine Basis $\{v_1, v_2, \cdots, v_{l'-1}, \bar{v}\}$ von $\mathfrak{B}(U)$ auf, so sind, weil die rationalen Vielfachen von $\bar{v}$ die einzigen Elemente von $\mathfrak{B}(U)$ sind, die in G homolog 0 sind, die Elemente $v_1, v_2, \cdots, v_{l'-1}$ in G linear unabhängig und können durch Elemente $v_{l'}, v_{l'+1}, \cdots, v_l$ zu einer Basis von $\mathfrak{B}(G)$ ergänzt werden. Wegen $d(\bar{v}) = m - 1$ gilt

$$d(v_{l'}) + d(v_{l'+1}) + \cdots + d(v_l) = d(G) - (d(U) - d(v)) = 2m - 1.$$

Nun kann aber, genau wie in a), unter den Elementen $v_{l'}, v_{l'+1}, \cdots, v_l$ keines vorkommen, dessen Dimension kleiner als m ist; folglich kann es nur ein solches Element geben, und es muss die Dimension $2m-1$ haben.

Das bedeutet: Es gibt in $\mathfrak{B}(U)$ ein Element $\bar{v}$ der Dimension $m-1$, dass nicht Null ist; und aus einer Basis von $\mathfrak{B}(U)$, in der dieses Element $\bar{v}$ als Basiselement auftritt, erhält man eine Basis von $\mathfrak{B}(G)$, indem man $\bar{v}$ ersetzt durch ein gewisses Element v_l der Dimension $2m-1$.

Aus $\bar{v} \sim 0$ in G folgt nach Kap. II, §2, Nr. 2, Korollar 1, dass auch $U \sim 0$ in G ist.

Damit ist die Behauptung b) von Nr. 1 vollständig bewiesen.

2. Die Gruppen A_n, B_n, C_n, D_n

Mit Hilfe des Satzes IV aus §1 lassen sich die Homologieringe der klassischen Gruppen A_n, B_n, C_n, D_n leicht bestimmen, auf Grund der Tatsache, dass

diesen Gruppen in natürlicher Weise Sphären als Wirkungsräume zugeordnet sind.

1. Die Gruppe A_n $(n = 1, 2, \cdots)$ ist die Gruppe der linearen, unitären, unimodularen Transformationen in $n + 1$ Variabeln, also die Gruppe der Transformationen

$$z_j' = \sum a_{jk} z_k$$

mit der Determinante $|a_{jk}| = 1$, die die Hermite'sche Form

$$z_1 \cdot \bar{z}_1 + z_2 \cdot \bar{z}_2 + \cdots + z_{n+1} \cdot \bar{z}_{n+1}$$

invariant lassen [23]. Wir setzen $z_j = x_j + iy_j$ (x_j, y_j reell); die Hermite'sche Form geht dabei über in

$$x_1^2 + y_1^2 + x_2^2 + y_2^2 + \cdots + x_{n+1}^2 + y_{n+1}^2 .$$

Die Transformationen von A_n lassen sich also auffassen als (reelle) orthogonale Transformationen eines Euklidischen R_{2n+2} mit den Koordinaten x_1, y_1, x_2, y_2, $\cdots$, x_{n+1}, y_{n+1}, die paarweise zu den komplexen Koordinaten $z_j = x_j + iy_j$ zusammengefasst sind. A_n ist also eine Gruppe von topologischen (sogar orthogonalen) Transformationen der $(2n + 1)$-dimensionalen Sphäre S_{2n+1}, die durch

$$x_1^2 + y_1^2 + x_2^2 + y_2^2 + \cdots + x_{n+1}^2 + y_{n+1}^2 = 1$$

gegeben ist. Man sieht leicht, dass diese Gruppe die S_{2n+1} transitiv transformiert. Es genügt zu beweisen, dass man den Punkt $(z_1 = 1, z_i = 0\ (i > 1))$ in einen beliebigen Punkt $(a_1, a_2, \cdots, a_{n+1})$ (mit $\sum a_j \cdot \bar{a}_j = 1$) überführen kann. Und das kann man, weil man immer unitäre, unimodulare Matrizen angeben kann, deren erste Spalte $(a_{11}, a_{21}, \cdots, a_{n+1,1})$ mit $(a_1, a_2, \cdots, a_{n+1})$ übereinstimmt. S_{2n+1} ist also ein Wirkungsraum von A_n (die Bedingung $T_p \cdot T_q = T_{p \cdot q}$ aus Kap. II, §2, Nr. 4 ist natürlich erfüllt). Die Isotropiegruppe ist die Untergruppe, die einen Punkt, etwa den Punkt $(z_1 = 1, z_i = 0\ (i > 1))$ festhält; das ist offenbar die Gruppe A_{n-1}.

$2n + 1$ ist ungerade; man erhält also nach Satz IV a) (§1): Der Homologiering $\mathfrak{R}(A_n)$ ist isomorph dem Ring $\mathfrak{R}(A_{n-1} \times S_{2n+1})$. Die Mannigfaltigkeit A_1 ist bekanntlich der Sphäre S_3 homöomorph. Damit erhält man durch Induktion [24]:

Der Homologiering $\mathfrak{R}(A_n)$ *ist isomorph dem Ring* $\mathfrak{R}(\Pi)$ *der Mannigfaltigkeit*

$$\Pi = S_3 \times S_5 \times S_7 \times \cdots \times S_{2n+1} ;$$

dabei bezeichnet S_i die i-dimensionale Sphäre.

Definiert man für einen beliebigen Komplex k das Poincaré'sche Polynom $P_k(t)$ durch:

$$P_k(t) = p_0 + p_1 \cdot t + p_2 \cdot t^2 + \cdots ,$$

wobei p_i die i-te Betti'sche Zahl von k und t eine Variable ist, und beachtet, dass

$$P_{k_1 \times k_2}(t) = P_{k_1}(t) \cdot P_{k_2}(t)$$

ist, so liest man ab:

Das Poincaré'sche Polynom von A_n ist

$$P_{A_n}(t) = (1 + t^3) \cdot (1 + t^5) \cdot (1 + t^7) \cdots (1 + t^{2n+1}).$$

2. Die Gruppe B_n ist die orthogonale Gruppe in $2n + 1$ Variabeln [23]. Das ist eine transitive Transformationsgruppe der Sphäre S_{2n}; also ist S_{2n} ein Wirkungsraum von B_n.

Die Gruppe D_n ist die orthogonale Gruppe in $2n$ Variabeln [23]. Das ist eine transitive Transformationsgruppe der Sphäre S_{2n-1}; also ist S_{2n-1} ein Wirkungsraum von D_n.

B_n ist definiert für $n = 1, 2, \cdots$; D_n wird bei der Aufzählung der einfachen Gruppen nur für $n \geqq 3$ betrachtet; für unsere Zwecke können wir aber die Fälle $n = 1, 2$ in die Definition von D_n einbeziehen. D_1 ist dem Kreis S_1 homöomorph.

Die Isotropiegruppe des Wirkungsraumes S_{2n-1} von D_n ist die Gruppe B_{n-1}; das sieht man, wenn man als Punkt der Sphäre $\sum x_i^2 = 1$, den man festhält, den Punkt $(1, 0, 0, \cdots, 0)$ wählt. $2n - 1$ ist ungerade; aus Satz IV a) folgt also:

Der Ring $\mathfrak{R}(D_n)$ ist isomorph dem Ring $\mathfrak{R}(B_{n-1} \times S_{2n-1})$. Dabei ist $n > 1$ vorausgesetzt; $\mathfrak{R}(D_1)$ ist isomorph $\mathfrak{R}(S_1)$.

Die Isotropiegruppe des Wirkungsraumes S_{2n} von B_n ist die Gruppe D_n; das sieht man wieder, wenn man den Punkt $(1, 0, \cdots, 0)$ festhält. $2n$ ist gerade; also kann man Satz IV b) anwenden. Danach hat man $\mathfrak{R}(D_n)$ als $\mathfrak{R}(\Pi \times S_{2n-1})$ darzustellen, und erhält dann $\mathfrak{R}(B_n)$ als $\mathfrak{R}(\Pi \times S_{4n-1})$. Wegen des eben erhaltenen Resultats über $\mathfrak{R}(D_n)$ kann man B_{n-1} für Π einsetzen; damit ist gezeigt:

Der Ring $\mathfrak{R}(B_n)$ ist isomorph dem Ring $\mathfrak{R}(B_{n-1} \times S_{4n-1})$, für $n > 1$. Für $n = 1$, folgt, nach der Bemerkung nach Satz IV, aus der Isomorphie von $\mathfrak{R}(D_1)$ mit $\mathfrak{R}(S_1)$, dass $\mathfrak{R}(B_1)$ mit $\mathfrak{R}(S_3)$ isomorph ist; das ergibt sich auch daraus, dass B_1 bekanntlich mit dem projektiven Raum P_3 homöomorph ist.

Durch Induktion erhält man jetzt sofort [24]:

Der Homologiering $\mathfrak{R}(B_n)$ ist isomorph dem Ring $\mathfrak{R}(\Pi)$ der Mannigfaltigkeit

$$\Pi = S_3 \times S_7 \times S_{11} \times \cdots \times S_{4n-1};$$

Der Homologiering $\mathfrak{R}(D_n)$ ist isomorph dem Ring $\mathfrak{R}(\Pi)$ der Mannigfaltigkeit

$$\Pi = S_3 \times S_7 \times S_{11} \times \cdots \times S_{4n-5} \times S_{2n-1} \qquad (\textit{für } n > 1);$$

$\mathfrak{R}(D_1)$ ist isomorph $\mathfrak{R}(S_1)$.

Für die Poincaré'schen Polynome entnimmt man daraus:

$P_{B_n}(t) = (1 + t^3)(1 + t^7) \cdots (1 + t^{4n-1})$;

$P_{D_n}(t) = (1 + t^3)(1 + t^7) \cdots (1 + t^{4n-5})(1 + t^{2n-1})$ *für* $n > 1$; $P_{D_1}(t) = 1 + t$.

3. Die Gruppe C_n ($n = 1, 2, \cdots$) ist die symplektische Gruppe [23, 25]: die Gruppe derjenigen unitären Transformationen eines $\bar{R}_{2n}$ mit den (komplexen) Koordinaten $(x_1, x_1', x_2, x_2', \cdots, x_n, x_n')$, die die schiefsymmetrische Bilinearform

$$(x_1y_1' - x_1'y_1) + (x_2y_2' - x_2'y_2) + \cdots + (x_ny_n' - x_n'y_n)$$

invariant lassen.

Den $\bar{R}_{2n}$ kann man, indem man die Koordinaten x_j, x_j' in Real- und Imaginärteil zerlegt, als einen reellen, Euklidischen, R_{4n} auffassen; die Transformationen von C_n sind dann wegen ihrer Unitarität in $\bar{R}_{2n}$ orthogonale Transformationen der $(4n - 1)$-dimensionalen Sphäre (etwa mit dem Radius 1) um den Nullpunkt des R_{4n}. Ist $(c_1, c_1', \cdots, c_n, c_n')$ ein beliebiger Punkt dieser Sphäre (ist also $c_1\bar{c}_1 + c_1'\bar{c}_1' + \cdots + c_n'\bar{c}_n' = 1$), so kann man eine Matrix aus C_n angeben, deren erste Spalte $(c_{11}, c_{11}', c_{21}, c_{21}', \cdots, c_{n1}, c_{n1}')$ mit $(c_1, c_1', \cdots, c_n, c_n')$ übereinstimmt [25]; das bedeutet: die S_{4n-1} wird transitiv transformiert, sie ist ein Wirkungsraum von C_n.

Die zugehörige Isotropiegruppe ist C_{n-1}: Man halte den Punkt $(1, 0, 0, \cdots, 0)$ (im $\bar{R}_{2n}$) fest; dann rechnet man (unter Beachtung der Invarianz der Bilinearform und der Unitarität) sofort nach, dass auch der Punkt $(0, 1, 0, \cdots, 0)$ des $\bar{R}_{2n}$ festbleibt. Beachtet man noch einmal die Unitarität, so erkennt man: Die Isotropiegruppe besteht aus den Transformationen von C_n, für die

$$x_1 \to x_1, \qquad x_1' \to x_1'$$

gilt, und das ist eben die Gruppe C_{n-1}.

$4n - 1$ ist ungerade. Nach Satz IV a) ist also $\mathfrak{R}(C_n)$ isomorph mit $\mathfrak{R}(C_{n-1} \times S_{4n-1})$. Beachtet man noch, dass die Gruppe C_1 mit der Gruppe A_1 identisch ist, so erhält man durch Induktion [24]:

Der Homologiering $\mathfrak{R}(C_n)$ *ist isomorph dem Ring* $\mathfrak{R}(\Pi)$ *der Mannigfaltigkeit*

$$\Pi = S_3 \times S_7 \times S_{11} \times \cdots \times S_{4n-1}$$

Daraus liest man ab: *Das Poincaré'sche Polynom von* C_n *ist*

$$P_{C_n}(t) = (1 + t^3)(1 + t^7) \cdots (1 + t^{4n-1}).$$

Kapitel IV

Homologieeigenschaften der Wirkungsräume

Nachdem wir im Kapitel III den speziellen Fall $W = S_m$ behandelt haben, wollen wir jetzt zur Untersuchung beliebiger Wirkungsräume übergehen. Die Untersuchung, die wir im Folgenden durchführen, ermöglicht es, wenn $U \sim 0$ in G ist, den Ring $\mathfrak{R}(W)$ durch $\mathfrak{R}(G)$ und $\mathfrak{R}(U)$ zu bestimmen.

1. Die Gruppe $P(\mathfrak{B}(G))$

1. Wie immer, bedeute G eine Gruppe, U eine (zusammenhängende) Untergruppe, W den zugehörigen Wirkungsraum G/U. Dann ist W eine orientierbare, geschlossene Mannigfaltigkeit der Dimension $d(W) = d(G) - d(U)$. Die Projektion P von G auf W bewirkt eine Abbildung P der Betti'schen Gruppe $\mathfrak{B}(G)$ in die Betti'sche Gruppe $\mathfrak{B}(W)$. Die Gruppe derjenigen Elemente von $\mathfrak{B}(W)$, die vermöge P Bilder von Elementen von $\mathfrak{B}(G)$ sind, heisse die Bildgruppe $P(\mathfrak{B}(G))$; ihre Struktur soll untersucht werden.

Wenn U ein Normalteiler von G ist, dann ist P die natürliche homomorphe Abbildung von G auf die Faktorgruppe G/U, hat also die in Kap. II, §1 beschriebenen Eigenschaften. Wir werden nun zeigen, dass auch dann, wenn U kein Normalteiler ist, die Abbildung P von diesem Typ ist; nur muss berücksichtigt werden, dass es in W im allgemeinen keine Multiplikation gibt.

2. Es sei also U eine beliebige (zusammenhängende) Untergruppe von G. Wir beweisen dann den

Satz V. *Der Kern der Abbildung P von $\mathfrak{B}(G)$ in $\mathfrak{B}(W)$ ist das Ideal in $\mathfrak{P}(G)$ (aufgefasst als Untergruppe von $\mathfrak{B}(G)$), das erzeugt wird von denjenigen minimalen Elementen von G, deren Bild bei P die Null ist.*

Die Restklassengruppe von $\mathfrak{B}(G)$ nach diesem Ideal (das Ideal als Untergruppe von $\mathfrak{B}(G)$ aufgefasst) wird also isomorph auf die Bildgruppe abgebildet. (Die Elemente dieser Restklassengruppe bilden zugleich den Restklassenring von $\mathfrak{P}(G)$ nach dem Ideal.)

Der Beweis geht aus von dem Aufspannsatz in der Form I (Kap. I, §3, Nr. 2) und wird folgendermassen geführt:

Die Gruppe $\mathfrak{B}(G)$ ist durch P homomorph in $\mathfrak{B}(W)$ abgebildet. $\mathfrak{B}^0(G)$ sei der Kern dieser Abbildung; $\mathfrak{B}(G)$ werde als direkte Summe $\mathfrak{B}^0(G) + \mathfrak{B}^1(G)$ dargestellt; $\mathfrak{B}^1(G)$ wird also isomorph abgebildet. $\{v_1, v_2, \cdots, v_m\}$ sei eine Basis von $\mathfrak{B}^1(G)$, und $\{v_{m+1}, v_{m+2}, \cdots, v_l\}$ eine von $\mathfrak{B}^0(G)$; dann ist $\{v_1, v_2, \cdots, v_l\}$ eine Basis von $\mathfrak{B}(G)$. Die Produkte

$$v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r} \qquad (i_1 < i_2 < \cdots < i_r)$$

bilden mit v_0 zusammen eine Basis $\langle v_i \rangle$ von $\mathfrak{B}(G)$ nach Satz I. Dann beweisen wir:

a) Ein solches Produkt wird durch P auf 0 gebildet, wenn wenigstens ein Faktor v_{i_k} zu $\mathfrak{B}^0(G)$ gehört, wenn also wenigstens ein Index $i_k > m$ ist;

b) diejenigen Produkte, in denen alle Faktoren v_{i_k} zu $\mathfrak{B}^1(G)$ gehören, in denen also alle Indizes $i_k \leqq m$ sind, werden durch P auf linear unabhängige Elemente von $\mathfrak{B}(W)$ abgebildet. Aus a) und b) folgt dann sofort der Satz V: Die in a) auftretenden Elemente bilden eine Basis des von $\mathfrak{B}^0(G)$ erzeugten Ideals von $\mathfrak{P}(G)$; dieses Ideal ist das in Satz V genannte. Aus a) folgt, dass das von $\mathfrak{B}^0(G)$ erzeugte Ideal durch P auf 0 abgebildet wird; aus b) entnimmt man, dass nur die Elemente dieses Ideals auf 0 abgebildet werden, womit Satz V bewiesen ist.

Man kann b) auch so formulieren: der von v_0 und v_1, v_2, $\cdots$, v_m erzeugte Teilring von $\mathfrak{B}(G)$ wird eineindeutig abgebildet; wegen a) wird er *auf* die Bildgruppe abgebildet. Bezeichnet S_i eine Sphäre mit $d(S_i) = d(v_i)$ für $i = 1, 2, \cdots, m$, so sieht man daraus: Die Gruppe $P(\mathfrak{B}(G))$ ist isomorph der Betti'schen Gruppe der Mannigfaltigkeit $S_1 \times S_2 \times S_3 \times \cdots \times S_m$.

Aus a) und b) kann man noch folgenden Schluss ziehen, der den Charakter der Abbildung P besser verstehen lässt: v_1', v_2', $\cdots$, v_r' seien irgendwelche Elemente von $\mathfrak{B}(G)$; das Bild $P(v_1' \circ v_2' \circ \cdots \circ v_r')$ von $v_1' \circ v_2' \circ \cdots \circ v_r'$ ist dann und nur dann ~ 0 in W, wenn die Elemente $\eta_i' = P(v_i')$ in W linear abhängig sind.

Sind nämlich die η_i' unabhängig, so kann man die v_i' in eine Basis von $\mathfrak{B}^1(G)$ aufnehmen, und nach b) ist $P(v_1' \circ v_2' \circ \cdots \circ v_r')$ nicht Null.

Sind aber die η_i' abhängig, so kann man durch eine lineare Transformation der v_i' erreichen, dass $\eta_r' = P(v_r')$ gleich Null ist; das Produkt $v_1' \circ v_2' \circ \cdots \circ v_r'$ ist nach Kap. I, §3, Nr. 9 invariant gegen eine solche Transformation, und aus $P(v_r') = 0$ folgt nach a), dass auch $P(v_1' \circ v_2' \circ \cdots \circ v_r') = 0$ ist.

Wir kommen jetzt zum Beweis unseres Satzes.

3. Zuerst beweisen wir den Teil a).

Wir definieren eine Abbildung T des topologischen Produktes $G \times W$ auf W, indem wir setzen:

$$T(p \times q') = T_p(q'),$$

wo p ein Punkt von G, und q' ein Punkt von W ist; T_p ist erklärt in Kap. II, §2, Nr. 4. Schreibt man W als Nebengruppenraum, und bedeutet q' die Nebengruppe qU, dann ist $T(p \times q')$ die Nebengruppe pqU.

Der Punkt $T(p \times q')$ heisst das Produkt von p und q' und wird auch mit pq' bezeichnet. (Wenn Produkte $c \cdot d$ auftreten, so wird es immer klar sein, ob es sich um das Gruppenprodukt in G handelt, wenn nämlich $c, d \in G$, oder um das eben definierte Produkt, wenn nämlich $c \in G$, $d \in W$; das gleiche gilt für das sofort zu definierende verallgemeinerte Pontrjagin'sche Produkt; es werden deshalb die alten Zeichen für die Multiplikation benützt.)

Ueber dieses Produkt sei bemerkt: $p \cdot q'$ ist bei festem p als Abbildung von W in sich aufzufassen, nämlich als $T_p(q')$; diese Abbildung ist topologisch und zur Identität homotop—man lasse p nach e, dem Einheitspunkt von G, wandern.

Bei festem q' dagegen ist $p \cdot q'$ eine Abbildung von G auf W, die zu der Projektion P homotop ist: man lasse q' in den die Nebengruppe U selbst darstellenden Punkt e' von W wandern; $p \cdot e'$ ist dann die Nebengruppe pU, als Punkt von W betrachtet, und das ist eben $P(p)$.

Die Eigenschaft des Produktes $p \cdot q'$, die wir brauchen, ist eine Funktionalgleichung für P; es gilt nämlich:

$$P(p_1 \cdot p_2) = p_1 \cdot P(p_2);$$

p_1, p_2 sind zwei beliebige Punkte von G; links steht die Projektion des Produktes $p_1 \cdot p_2$; rechts steht das Produkt des Punktes p_1 von G mit dem Punkte

$P(p_2)$ von W, also $T(p_1 \times P(p_2))$; beide Seiten der Gleichung bedeuten die Nebengruppe $p_1 \cdot p_2 \cdot U$, als Punkt von W.

Die Abbildung T bewirkt eine Abbildung T der Homologieklassen. Wir definieren nun, wie ja nahe liegt, das "verallgemeinerte Pontrjagin'sche Produkt $x \circ y'$" eines Elementes x von $\mathfrak{B}(G)$ mit einem Element y' von $\mathfrak{B}(W)$, indem wir setzen:

$$x \circ y' = T(x \times y');$$

das ist wieder ein Element von $\mathfrak{B}(W)$.

Diese Multiplikation ist offenbar distributiv mit der Addition verknüpft. Es gilt

$$x \circ 0' = 0 \circ y' = 0',$$

wo 0, bzw. 0', das Nullelement von $\mathfrak{B}(G)$, bzw. $\mathfrak{B}(W)$, ist, und x, y' beliebig sind.

Die Funktionalgleichung für P gibt für die Homologieklassen die entsprechende Funktionalgleichung, nämlich:

$$P(x_1 \circ x_2) = x_1 \circ P(x_2),$$

wenn x_1, x_2 zwei beliebige Elemente von $\mathfrak{B}(G)$ sind. Das ergibt sich so: F sei die bekannte Abbildung von $G \times G$ in G; mit S bezeichne man die Abbildung von $G \times G$ in $G \times W$, die durch $S(p_1 \times p_2) = p_1 \times P(p_2)$ gegeben ist. Dann ist wegen der Funktionalgleichung die Abbildung PF von $G \times G$ in W gleich der Abbildung TS. Nun ist $PF(x_1 \times x_2) = P(x_1 \circ x_2)$; andererseits ist $S(x_1 \times x_2) = x_1 \times P(x_2)$ (vgl. Kap. I, §1, Nr. 8), also $TS(x_1 \times x_2) = x_1 \circ P(x_2)$.

Jetzt folgt sofort die Behauptung a). $v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r}$ sei ein Element der Basis $\langle v_i \rangle$, und es sei $P(v_{i_k}) = 0'$, $k \leqq r$. Wegen der Antikommutativität der v_i darf man $k = r$ annehmen, und kann also das Element in der Form $w \circ v_{i_r}$ mit $w = v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_{r-1}}$ schreiben. Dann ist

$$P(w \circ v_{i_r}) = w \circ P(v_{i_r}) = w \circ 0' = 0'.$$

4. Jetzt beweisen wir die Behauptung b); wir benützen den Umkehrungshomomorphismus φ der Abbildung P.

Als Basis in $\mathfrak{B}(G)$ benützen wir die von v_0 und den Elementen

$$v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r} \qquad (i_1 < i_2 < \cdots < i_r)$$

gebildete Basis $\langle v_i \rangle$. Nach dem Aufspannsatz (vgl. Kap. I, §3, Nr. 9) ist die duale Basis die Basis $[z_i]$, die erzeugt wird von den zu den v_i dualen Elementen z_i.

Die Bilder $\eta_i = P(v_i)$ der Elemente $v_1, v_2, \cdots, v_m$ sind nach Voraussetzung in W linear unabhängig. Sie werden in eine Betti'sche Basis von $\mathfrak{B}(W)$ aufgenommen; ζ_i $(i = 1, 2, \cdots, m)$ seien die in der dualen Basis zu den η_i dualen Elemente.

Wir betrachten $\varphi(\zeta_i)$. Wir behaupten: stellt man $\varphi(\zeta_i)$ in der Basis $[z_i]$ dar, dann ist z_i das einzige maximale Basiselement, das in $\varphi(\zeta_i)$ auftritt, und zwar hat es den Koeffizienten 1. Wir behaupten also:

$$(\zeta) \qquad \varphi(\zeta_i) = z_i + \sum a_h x_h,$$

wo die x_h zusammengesetzte Elemente der Basis $[z_i]$ und die a_h Koeffizienten sind.

Zum Beweis betrachten wir das Produkt $\varphi(\zeta_i)\cdot v_k$. Die Funktionalgleichung (Kap. I, §1, Nr. 5) für φ liefert

$$P(\varphi(\zeta_i)\cdot v_k) = \zeta_i\cdot P(v_k) = \zeta_i\cdot\eta_k\,, \qquad \text{bzw. } = \zeta_i\cdot 0'$$

also

$$P(\varphi(\zeta_i)\cdot v_k) = \delta_{ik}\cdot\eta_0 \qquad (i = 1, \cdots, m;\ k = 1, \cdots, l),$$

wo $\eta_0 = P(v_0)$ der einfach gezählte Punkt von W ist; denn weil η_k minimal ist, ist $\zeta_k\cdot\eta_k = \eta_0$ der einzige von Null verschiedene Schnitt von η_k, für $k = 1, 2, \cdots, m$; für $k > m$ ist $P(v_k) = 0'$. Also ist

$$\varphi(\zeta_i)\cdot v_k = \delta_{ik}v_0\,;$$

daraus folgt auf Grund der Rechenregeln in $\mathfrak{R}(G)$ sofort die Behauptung (ζ)— man beachte etwa die Gleichungen $z_i\cdot v_k = \delta_{ik}v_0$ ($i, k = 1, \cdots, l$) (Kap. I, §3, Nr. 8).

Wegen der Multiplikativität von φ hat man dann:

$$\varphi(\zeta_{i_1}\cdot\zeta_{i_2}\cdot\ \cdots\ \cdot\zeta_{i_r}) = z_{i_1}\cdot z_{i_2}\cdot\ \cdots\ \cdot z_{i_r} + \sum y_h\,,$$

wo jedes y_h ein Produkt von mindestens $r+1$ Faktoren z_i ist.

Aus dieser Darstellung leitet man die folgenden Formeln ab (vgl. Kap. I, §3, Nr. 9):

1) $$\varphi(\zeta_{i_1}\cdot\zeta_{i_2}\cdot\ \cdots\ \cdot\zeta_{i_r})\cdot(v_{i_1}\circ v_{i_2}\circ\ \cdots\ \circ v_{i_r}) = \pm v_0\,,$$

2) $$\varphi(\zeta_{i_1}\cdot\zeta_{i_2}\cdot\ \cdots\ \cdot\zeta_{i_r})\cdot(v_{j_1}\circ v_{j_2}\circ\ \cdots\ \circ v_{j_r}) = 0,$$

wenn die Indexmengen $\{i_1, i, \cdots, i_r\}$ und $\{j_1, j_2, \cdots, j_r\}$ voneinander verschieden sind,

3) $$\varphi(\zeta_{i_1}\cdot\zeta_{i_2}\cdot\ \cdots\ \cdot\zeta_{i_r})\cdot(v_{k_1}\circ v_{k_2}\circ\ \cdots\ \circ v_{k_s}) = 0, \qquad \text{wenn } s < r.$$

5. Daraus folgt nun leicht die Behauptung b) in der Formulierung: der von $v_0, v_1, v_2, \cdots, v_m$ erzeugte Teilring von $\mathfrak{P}(G)$ wird eineindeutig abgebildet.

Es sei nämlich w ein Element dieses Teilringes, also eine Linearkombination $\sum a_i w_i$ (mit Koeffizienten a_i), in der die w_i solche Elemente[6]

$$v_{i_1}\circ v_{i_2}\circ\ \cdots\ \circ v_{i_r} \qquad (i_1 < i_2 < \cdots < i_r),$$

bedeuten, in denen $i_k \leqq m$ gilt; die Koeffizienten a_i seien von Null verschieden. $w_1 = v_{h_1}\circ v_{h_2}\circ\ \cdots\ \circ v_{h_t}$ sei eines dieser Elemente mit maximaler Anzahl von Faktoren v_i. Dann ist nach Nr. 4, 1),

$$\varphi(\zeta_{h_1}\cdot\zeta_{h_2}\cdot\ \cdots\ \cdot\zeta_{h_t})\cdot w_1 = \pm v_0\,,$$

und nach Nr. 4, 2) und 3),

$$\varphi(\zeta_{h_1}\cdot\zeta_{h_2}\cdot\ \cdots\ \cdot\zeta_{h_t})\cdot w_i = 0 \qquad \text{für } i > 1:$$

also ist schliesslich

$$\varphi(\zeta_{h_1}\cdot\zeta_{h_2}\cdot\ \cdots\ \cdot\zeta_{h_t})\cdot w = \pm a_1 v_0\,.$$

---

[6] Wir dürfen $w_i \neq v_0$ annehmen.

Nach der Funktionalgleichung für φ folgt hieraus

$$(\zeta_{h_1}\cdot\zeta_{h_2}\cdot\ \cdots\ \cdot\zeta_{h_t})\cdot P(w) = \pm a_1\eta_0\,. \tag{1}$$

Also ist $P(w) \neq 0$, q.e.d.

Es sei noch bemerkt, dass das Produkt aller ζ_i, das Element $\zeta_1\cdot\zeta_2\cdot\ \cdots\ \cdot\zeta_m$, von Null verschieden ist: man setze für w das Produkt $v_1\circ v_2\circ\ \cdots\ \circ v_m$ ein.

2. $U \sim 0$

1. Für den Fall, dass $U \sim 0$ in G ist, lässt sich jetzt eine sehr präzise Aussage herleiten. Wir beweisen nämlich den folgenden Satz:

Satz VI. *Die Untergruppe U der Gruppe G sei nicht homolog 0 in G; der zugehörige Wirkungsraum sei $G/U = W$; dann gilt:*

a) *der Ring $\mathfrak{R}(W)$ ist isomorph dem Ring eines topologischen Produktes von Sphären ungerader Dimensionen,*

b) *der Ring $\mathfrak{R}(G)$ ist isomorph dem Ring $\mathfrak{R}(U \times W)$ des topologischen Produktes $U \times W$.*

Danach ist $\mathfrak{R}(W)$ durch $\mathfrak{R}(G)$ und $\mathfrak{R}(U)$ bestimmt: $l_i(W)$, die Zahl der i-dimensionalen Sphären in dem $\mathfrak{R}(W)$ darstellenden Sphärenprodukt, ist gegeben durch $l_i(G) - l_i(U)$.

Zum Beweis zeigen wir erstens, dass die Projektion P eine Abbildung von $\mathfrak{R}(G)$ *auf* $\mathfrak{R}(W)$ liefert; nach §1 ist dann $\mathfrak{R}(W)$ eineindeutiges Bild eines gewissen Teilringes des Pontrjagin'schen Ringes $\mathfrak{P}(G)$. Zweitens bestimmen wir diesen Teilring; er ergibt sich als isomorph dem Restklassenring des von den Elementen von $\mathfrak{B}(U)$ erzeugten Ideals von $\mathfrak{P}(G)$. Daraus folgt dann leicht der Satz.

2. U ist ~ 0 in G; die Homologieklasse von G, in der U liegt, bezeichnen wir auch mit U. Unter φ verstehen wir, wir in §1, den Umkehrungshomomorphismus der Projektion P. Dann gilt:

$$\varphi(\eta_0) = U.$$

Dabei ist η_0 der einfach gezählte Punkt in W.

Es gilt nämlich allgemein für die Projektion eines gefaserten Raumes auf den Faserraum: Der Urbildzyklus (das φ-Bild) des Punktes ist die Faser. Das folgt z.B. daraus, dass man die Umkehrung φ einer solchen Abbildung P folgendermassen erklären kann: Der Faserraum B sei hinreichend fein trianguliert. Für ein Simplex x_i von B definiert man $\varphi(x_i)$ als die, in geeigneter Weise als Komplex aufgefasste, Menge $P^{-1}(x_i)$, die fasertreu homöomorph mit $x_i \times F$ (F = Faser) ist. Für Komplexe, Zyklen usw. erhält man φ dann durch Addition.

Wegen der Voraussetzung $U \sim 0$ in G ist also $\varphi(\eta_0) \neq 0$. Nach dem Hilfssatz aus Kap. I, §1, Nr. 6 ist dann P eine Abbildung von $\mathfrak{R}(G)$ *auf* $\mathfrak{R}(W)$.

Zieht man jetzt den Satz V heran, so sieht man, dass die Gruppe $\mathfrak{B}(W)$ bestimmt ist, wenn man die Gruppe $\mathfrak{B}^0(G)$, d.h. die Gruppe derjenigen mini-

malen Elemente von $\mathfrak{R}(G)$, deren P-Bild in W homolog 0 ist kennt. Diese Gruppe soll jetzt bestimmt werden.

3. Weil U $\sim$ 0 ist, ist kein Element (ausser 0) von $\mathfrak{B}(U)$ homolog 0 in G (Kap. II, §2, Nr. 2, Korollar 1), und man kann $\mathfrak{B}(U)$ als Untergruppe von $\mathfrak{B}(G)$ auffassen. Die Elemente von $\mathfrak{B}(U)$ werden bei P sämtlich auf 0 abgebildet: die ganze Untergruppe U wird ja auf einen Punkt abgebildet. Wir behaupten: mit den Elementen von $\mathfrak{B}(U)$ sind die Elemente von $\mathfrak{B}(G)$, deren P-Bild die Null ist, erschöpft, oder in der Bezeichnung von §1, Nr. 2: es ist $\mathfrak{B}^0(G) = \mathfrak{B}(U)$.

Zum Beweis setzen wir $l(G) = l$ und $l(U) = l'$. Die Elemente $v_1, v_2, \cdots, v_{l'}$ mögen eine Basis von $\mathfrak{B}(U)$ bilden; sie werden aufgenommen in eine Basis $\{v_1, v_2, \cdots, v_l\}$ von $\mathfrak{B}(G)$. Dann gilt (nach dem Aufspannsatz, angewendet auf die Gruppe U; vgl. Kap. I, §3, Nr. 7)

$$v_1 \circ v_2 \circ \cdots \circ v_{l'} = U$$

als Homologie in U (dann bedeutet U die Eins von $\mathfrak{R}(U)$); also gilt die Gleichung auch als Homologie in G. Wir setzen noch

$$v_{l'+1} \circ v_{l'+2} \circ \cdots \circ v_l = V.$$

Dann gilt, wieder nach den Aufspannsatz, die Gleichung

$$U \circ V = 1,$$

wo 1 die Eins von $\mathfrak{R}(G)$ bedeutet; daher gilt (nach Kap. I, §3, Nr. 10) für den Schnitt $U \cdot V$:

$$U \cdot V = v_0 .$$

Uebt man auf diese Gleichung die Abbildung P aus, und beachtet, dass $U = \varphi(\eta_0)$ ist, so erhält man wegen der Funktionalgleichung für φ die Gleichung:

$$\eta_0 \cdot P(V) = \eta_0 .$$

Bezeichnet man die Eins von $\mathfrak{R}(W)$ mit $\bar{1}$, so heisst das:

$$P(V) = \bar{1}.$$

Es ist also $P(v_{l'+1} \circ v_{l'+2} \circ \cdots \circ v_l) \neq 0$; daraus folgt nach der letzten Bemerkung in §1, Nr. 2, dass die Bilder $\eta_i = P(v_i)$ der Elemente $v_{l'+1}, v_{l'+2}, \cdots, v_l$ in W linear unabhängig sind. Das bedeutet aber: die Restklassengruppe $\mathfrak{B}(G) - \mathfrak{B}(U)$ wird durch P isomorph abgebildet.

Daher ist $\mathfrak{B}(U)$ der Kern der Abbildung P von $\mathfrak{B}(G)$; die Behauptung $\mathfrak{B}^0(G) = \mathfrak{B}(U)$ ist bewiesen.

Die von $v_{l'+1}, v_{l'+2}, \cdots, v_l$ erzeugte Untergruppe von $\mathfrak{B}(G)$ ist die in §1, Nr. 2 mit $\mathfrak{B}^1(G)$ bezeichnete Gruppe; der von dem Punkt v_0 und den Elementen dieser Gruppe erzeugte Teilring von $\mathfrak{B}(G)$ wird nach §1, Nr. 2 eineindeutig auf die Bildgruppe abgebildet. Da jetzt nach Nr. 2 die Bildgruppe mit der ganzen

Gruppe $\mathfrak{B}(W)$ zusammenfällt, ist also $\mathfrak{B}(W)$ eineindeutiges Bild des von den $v_{l'+1}$, $v_{l'+2}$, $\cdots$, v_l und v_0 erzeugten Teilringes von $\mathfrak{P}(G)$.

Dieser Teilring enthält nun genau $2^{l-l'}$ linear unabhängige Elemente, z.B. den Punkt v_0 und die Produkte

$$v_{i_1} \circ v_{i_2} \circ \cdots \circ v_{i_r} \qquad (l' + 1 \leqq i_k \leqq l,\ i_1 < i_2 < \cdots < i_r).$$

Also enthält auch $\mathfrak{B}(W)$ genau $2^{l-l'}$ linear unabhängige Elemente.

4. Wir betrachten die zu den $\eta_i = P(v_i)$ dualen Elemente ζ_i von $\mathfrak{B}(W)$ $(i = l' + 1, \cdots, l)$. Weil $P(V) = P(v_{l'+1} \circ v_{l'+2} \circ \cdots \circ v_l) = \bar{1}$ die Dimension $d(W)$ hat, ist der Schnitt $\zeta_{l'+1} \cdot \zeta_{l'+2} \cdot \cdots \cdot \zeta_l$ nulldimensional. Aus der Formel (1) in §1, Nr. 5, entnimmt man leicht, dass dieser Schnitt gleich $\pm\eta_0$, also von Null verschieden ist; man setze dazu für das dort mit w bezeichnete Element das Produkt $V = v_{l'+1} \circ v_{l'+2} \circ \cdots \circ v_l$ ein und beachte, dass $P(V) = \bar{1}$ ist. Ausserdem sind die ζ_i antikommutativ in Bezug auf den Schnitt:

$$\zeta_i \cdot \zeta_k = -\zeta_k \cdot \zeta_i \quad \text{und} \quad \zeta_i \cdot \zeta_i = 0,$$

denn es ist $\delta(\zeta_i) = d(\eta_i) = d(v_i)$ ungerade.

Daraus folgert man (vgl. die Betrachtung in Kap. I, §3, Nr. 7), dass die Produkte

$$\zeta_{i_1} \cdot \zeta_{i_2} \cdot \cdots \cdot \zeta_{i_r} \qquad (l' + 1 \leqq i_k \leqq l,\ i_1 < i_2 < \cdots < i_r)$$

linear unabhängige Elemente in $\mathfrak{B}(W)$ sind; mit $\bar{1}$ zusammen sind das genau $2^{l-l'}$ Elemente; die ζ_i bilden also mit $\bar{1}$ zusammen ein irreduzibles Erzeugendensystem von $\mathfrak{R}(W)$. Diese Eigenschaften der ζ_i sind es aber gerade, die $\mathfrak{R}(W)$ als Ring eines Produktes von ungerade-dimensionalen Sphären charakterisieren. Damit ist VI a) bewiesen. Den "Sphären selbst" entsprechen die Elemente η_i.

5. Den Isomorphismus von $\mathfrak{R}(G)$ und $\mathfrak{R}(U \times W)$ stellt man jetzt folgendermassen her:

1 sei die Eins von $\mathfrak{R}(G)$, $1'$ die von $\mathfrak{R}(U)$, $\bar{1}$ die von $\mathfrak{R}(W)$.

z_i seien die zu den v_i dualen Elemente von $\mathfrak{B}(G)$ $(i = 1, \cdots, l)$; sie bilden mit 1 zusammen ein irreduzibles Erzeugendensystem von $\mathfrak{R}(G)$ (vgl. Kap. I, §2, Nr. 4).

z_j' seien die zu den v_j dualen Elemente von $\mathfrak{B}(U)$ $(j = 1, \cdots, l')$; sie bilden mit $1'$ zusammen ein irreduzibles Erzeugendensystem von $\mathfrak{R}(U)$.

ζ_k seien die oben eingeführten Elemente von $\mathfrak{B}(W)$ $(k = l' + 1, \cdots, l)$; sie bilden mit $\bar{1}$ zusammen ein irreduzibles Erzeugendensystem von $\mathfrak{R}(W)$.

Die Elemente $(z_j' \times \bar{1})$ und $(1' \times \zeta_k)$ bilden dann mit dem Element $(1' \times \bar{1})$ zusammen ein irreduzibles Erzeugendensystem von $\mathfrak{R}(U \times W)$.

Alle diese Systeme sind antikommutativ; und das Produkt aller Elemente jeden Systems ist der einfach gezählte Punkt des betreffenden Ringes.

Jetzt ordnet man zu:

Den Elementen z_i $(i = 1, \cdots, l')$ die Elemente $(z_j' \times \bar{1})$,

den Elementen z_i $(i = l' + 1, \cdots, l)$ die Elemente $(1' \times \zeta_k)$,
dem Element 1 das Element $1' \times \bar{1}$.

Das ergibt offenbar einen Isomorphismus zwischen $\mathfrak{R}(G)$ und $\mathfrak{R}(U \times W)$ der gewünschten Art; damit ist auch VI b) bewiesen.

ZÜRICH, SWITZERLAND

LITERATURVERZEICHNIS

1. L. Pontrjagin: *Topological Groups* (Princeton 1939), Def. 38, S. 181 u. S. 83, D).
2. E. Cartan: *La Topologie des . . . Groupes de Lie* [L'Enseignement Math. 35 (1936); dasselbe auch: Actualités Scient. et Industr. 358].
3. L. Pontrjagin: *On Homologies in compact Lie groups* [Recueil math. de Moscou, t. 6 (48), 3 (1939)].
4. R. Brauer: C. R. 201 (1935), S. 419.
5. Ch. Ehresmann: C. R. 208 (1939), S. 321 u. S. 1263.
6. A. a. O. [2], S. 199.
7. H. Hopf: *Ueber die Topologie der Gruppen-Mannigfaltigkeiten und ihrer Verallgemeinerungen* [these Annals (2) 42 (1941)], Satz I.
8. A. a. O. [2], S. 198. Diese Vermutung ist auch ausgesprochen bei Hopf, a. a. O. [7], Nr. 37.
9. E. Cartan: *La theorie des groupes finis et continus et l'analysis situs* [Memorial Sc. Math. XLII], Nr. 17.
10. J. Feldbau: C. R. 208 (1939), S. 1621.
11. S. Lefschetz: *Topology* (New York 1930).
12. H. Hopf: *Zur Algebra der Abbildungen von Mannigfaltigkeiten* [Journal f. d. r. u. a. Math. 163 (1930)].
    H. Freudenthal: *Zum Hopf'schen Umkehrhomomorphismus* [Annals of Math. (2) 38 (1937)].
    A. Komatu: *Über die Ringdualität eines Kompaktums* [Tôhoku Math. J. 43 (1937)].
    H. Whitney: *On Products in a complex* [Annals of Math. (2) 39 (1938)], Th. 6.
13. Alexandroff-Hopf: *Topologie I* (Berlin 1935), Kap. VII, §3.
14. A. a. O. [11].
15. A. a. O. [2], S. 198.
16. A. a. O. [7], Nr. 37.
17. A. a. O. [1], Th. 50, insbesondere S. 200, Formel (20).
18. S. S. Cairns: *On the triangulation of regular loci* [Annals of Math. (2) 35 (1934)].
    H. Freudenthal: *Die Triangulation der differenzierbaren Mannigfaltigkeiten* [Proceedings Amsterdam XLII (1939)].
19. A. a. O. [10], Th. A.
20. A. a. O. [13], S. 128 u. 149.
21. A. a. O. [13], S. 334.
22. A. a. O. [13], Kap. VII, §2.
23. A. a. O. [3], S. 397 f.
24. A. a. O. [3], Th. 2 u. 3.
25. H. Weyl: *The classical groups* (Princeton 1940), S. 163 u. S. 172.

ANNALS OF MATHEMATICS
Vol. 42, No. 5, December, 1941

INTRINSIC TOPOLOGY AND COMPLETION OF BOOLEAN RINGS

BY HENRY LÖWIG

(Received October 29, 1940)

NOTATION

If $\mathfrak{B}$ is a subclass of a lattice A we denote the meet (cross-cut, intersection, product) of the elements of $\mathfrak{B}$ with respect to A, i. e. the maximal element of A, contained in all elements of $\mathfrak{B}$, its existence supposed, by $\prod_{x \in \mathfrak{B}}^{(A)} x$, and the join (union, conjunction, lattice-sum) of the elements of $\mathfrak{B}$ with respect to A, i.e. the minimal element of A, containing all elements of $\mathfrak{B}$, its existence supposed, by $\sum_{x \in \mathfrak{B}}^{(A)} x$. If $\mathfrak{B}$ consists only of two elements, a and b, we write also $a \wedge_A b$ instead of $\prod_{x \in \mathfrak{B}}^{(A)} x$ and $a \vee_A b$ instead of $\sum_{x \in \mathfrak{B}}^{(A)} x$. If A is especially a Boolean ring we denote the ring-sum (the symmetric difference) of two elements a and b with respect to A by $a \triangle_A b$. We omit the superscripts of the signs $\prod^{(A)}$ and $\sum^{(A)}$, and the subscript of the sign $\vee_A$ if there is no doubt to which lattice or Boolean ring they refer. Under the same condition we write $a \cdot b$ or ab instead of $a \wedge_A b$, and $a + b$ instead of $a \triangle_A b$. We call a subclass $\mathfrak{B}$ of a lattice A *bounded* if there exists at least one couple a, b of elements of A such that

$$a < x < b$$

for all elements x of $\mathfrak{B}$. The expressions of "σ-lattice" and "complete lattice" are used in the same sense as in the paper (VIII). (See especially (VIII), p. 795, 4th passage.) We call a Boolean ring complete if it is a complete lattice, and a Boolean σ-ring if it is a σ-lattice. If a subring (an ideal) of a Boolean ring is a σ-lattice we call it a σ-subring (a σ-ideal). The expression of "dual ideal" is used in the same sense as in (VI). The words "A Boolean ring A is given, and the small Latin letters ... denote elements of A" will be frequently suppressed. If we consider a sequence of elements a_n of a lattice we shall mostly suppress the supposition that the index n runs through all natural numbers. Otherwise we shall use the terminology and denotation of M. H. Stone's paper (V). The results of this paper will be used without any further reference.

1. Introduction

A sequential topology of a lattice can be defined in the following way:

If a sequence of elements a_n of a lattice has the property that all products $\prod_{k=n}^{\infty} a_k$ $(n = 1, 2, 3, \cdots)$, all joins $\sum_{k=n}^{\infty} a_k$ $(n = 1, 2, 3, \cdots)$, the join $\sum_{n=1}^{\infty} \prod_{k=n}^{\infty} a_k$, and the product $\prod_{n=1}^{\infty} \sum_{k=n}^{\infty} a_k$ exist, and

$$\sum_{n=1}^{\infty} \prod_{k=n}^{\infty} a_k = \prod_{n=1}^{\infty} \sum_{k=n}^{\infty} a_k = a,$$

the sequence a_n is said to converge towards a or to have the limit a. (See for instance (II), p. 453.)

The lattices considered in L. V. Kantorovitch's paper (VII) have even the property that $\prod_{k=n}^{\infty} a_k$, $\sum_{k=n}^{\infty} a_k$ $(n = 1, 2, 3, \cdots)$, $\sum_{n=1}^{\infty} \prod_{k=n}^{\infty} a_k$, and $\prod_{n=1}^{\infty} \sum_{k=n}^{\infty} a_k$ exist whenever the sequence a_n is bounded. Thus Kantorovitch may say a sequence a_n to converge if it is bounded, and

$$\sum_{n=1}^{\infty} \prod_{k=n}^{\infty} a_k = \prod_{n=1}^{\infty} \sum_{k=n}^{\infty} a_k .$$

In this paper the above convergence-definition will be considerably generalized for the case that the given lattice is a *Boolean ring*. We shall define the convergence of a sequence of elements a_n of a Boolean ring without supposing that $\prod_{k=n}^{\infty} a_k$, $\sum_{k=n}^{\infty} a_k$ $(n = 1, 2, 3, \cdots)$, $\sum_{n=1}^{\infty} \prod_{k=n}^{\infty} a_k$, and $\prod_{n=1}^{\infty} \sum_{k=n}^{\infty} a_k$ exist; we shall not even suppose that the sequence a_n is bounded.

The formulation of the convergence-definition will be the topic of §2 of this paper. In §3 we shall show that in the case that $\prod_{k=n}^{\infty} a_k$ and $\sum_{k=n}^{\infty} a_k$ $(n = 1, 2, 3, \cdots)$ exist, our convergence-definition coincides with the definition given in the introduction. In §4 we shall introduce the important notion of an *invariant* subring $\mathfrak{a}$ of a given Boolean ring A. We shall call $\mathfrak{a}$ an invariant subring of A, or A an invariant extension of $\mathfrak{a}$, if every infinite join of elements of $\mathfrak{a}$, whenever it exists with respect to $\mathfrak{a}$, exists also with respect to A, and has the same value. In §5 we shall define the notion of a *join-extension* A of a Boolean ring $\mathfrak{a}$. A will be said to be a join-extension of $\mathfrak{a}$ if every element of A is a join of elements of $\mathfrak{a}$ with respect to A. We shall see that any join-extension is an invariant extension. The notions of "invariant extension" and "join-extension" will play an important rôle in the following paragraphs. In §6 we shall show that the finite fundamental operations of a Boolean ring are continuous under the introduced sequential topology; moreover we shall deduce a number of other relations. In §7 we shall say a few words about the closure-topology determined by the introduced sequential topology. In §8 we shall state how far the theorems on simple sequences in a Boolean ring obtained hitherto can be extended to double sequences. In §9 we shall, following D. van Dantzig's paper (I), define the notion of a fundamental sequence of elements of a Boolean ring, and explore how far D. van Dantzig's results can be applied to our case. In §10 we shall show that the sequential topology determined by the closure-topology mentioned in §7 is in general not identical with the original sequential topology. I have no further entered into the examination of this derivative sequential topology, but leave it to further investigations.

It is very probable that the theory developed in this paper can be easily extended to the lattices which Fritz Klein in his paper (IV) calls "Boole-Schrödersche Verbände." I shall deal with this matter in a later paper.

The extension of this theory to arbitrary lattices might be more complicated. It is for instance not true that the finite fundamental operations of an arbitrary lattice are continuous with respect to the sequential topology defined in this

introduction. Let a partially ordered set A consist of an enumerated sequence of elements a_n, of another such sequence b_n, and of a further element e, between which there are defined the following ordering relations:

$$a_m < a_n, \qquad b_m < b_n \qquad \text{for } m < n;$$

$$a_m < b_n \qquad \text{for } m \leqq n;$$

$$a_m < e, \qquad b_m < e.$$

$$(m, n = 1, 2, 3, \cdots .)$$

Then A is obviously a complete distributive lattice. ADDED IN PROOF: A is isomorphic with Mac Neille's lattice [4.7, 4.12, 4.7]. See H. M. Mac Neille, *Partially ordered sets*, Trans. Am. Math. Soc., vol. 42 (1937), pp. 416–460, especially Theorem 4.15. If $m > n$, then

$$a_m b_n = a_n,$$

$$a_m \vee b_n = b_m.$$

Obviously we have, by the convergence-definition given in this introduction

$$\lim_{n\to\infty} a_n = e;$$

on the other hand we have for $n = 1, 2, 3, \cdots$

$$a_n b_1 = a_1,$$

whence

$$\lim_{n\to\infty} a_n b_1 = a_1,$$

while

$$e b_1 = b_1.$$

2. The Sequential Topology $\tau(A)$

DEFINITION 1. *By a subelement of a sequence of elements a_n of a Boolean ring A, we mean an element u of A of the property that, if we suitably choose the natural number k, we have for $n \geqq k$*

$$u < a_n.$$

It is clear that the property of an element u of A to be a subelement of the sequence a_n is preserved if we replace a finite number of its terms by other elements, if we add or omit a finite number of terms, or if we change the succession of the terms (not necessarily only of a finite number). The same holds for the concepts of superelement, interelement, lower limit, upper limit, and limit, which we shall define more fully below.

Any sequence of elements of a Boolean ring has at least one subelement, namely the zero-element.

THEOREM 1. *The subelements of a sequence of elements a_n of a Boolean ring A constitute an ideal in A.*

THEOREM 2. *If u is a subelement of a sequence a_n, and b is an arbitrary element, then ub is a subelement of the sequence a_nb, and to every subelement u^* of the sequence a_nb there exists a subelement u of the sequence a_n such that $u^* = ub$.*

DEFINITION 2. *By a superelement of a sequence a_n, we mean an element o of the property that, if we suitably choose the natural number k, we have for $n \geqq k$*

$$o > a_n .$$

It is clear that a sequence of elements of a Boolean ring A is bounded in A if and only if it has at least one superelement in A. If A has a unit every sequence of elements of A is bounded.

THEOREM 3. *The superelements of a bounded sequence of elements of a Boolean ring A constitute a dual ideal in A.* (This dual ideal is a Boole-Schröder lattice in the sense of the paper (IV).)

THEOREM 4. *If o is a superelement of a bounded sequence a_n, and b an arbitrary element, then $o \vee b$ is a superelement of the sequence $a_n \vee b$, and to every superelement o^* of the sequence $a_n \vee b$ there exists a superelement o of the sequence a_n such that $o^* = o \vee b$.*

DEFINITION 3. *Let a_n be a bounded sequence of elements of a Boolean ring A, and also a an element of A. In this case we call a an interelement of the sequence a_n with respect to A, if*

$$u < a < o$$

whenever u and o are elements of A, and u is a subelement, and o a superelement of the sequence a_n.

DEFINITION 4. *Let a_n be an unbounded sequence of elements of a Boolean ring A, and also a an element of A. In this case we call a an interelement of the sequence a_n with respect to A if ab is, according to Definition 3, an interelement of the bounded sequence a_nb with respect to A for every element b of A.*

Instead of saying "a is an interelement of the sequence a_n with respect to A" we may say more simply "a is an interelement of the sequence a_n" if there is no doubt to which Boolean ring the term "interelement" is referred.

THEOREM 5. *If a_n and a are elements of a Boolean ring A, then a is an interelement of the sequence a_n with respect to A if and only if ab is an interelement of the sequence a_nb with respect to A for every element b of A.*

PROOF. As for the case of an unbounded sequence Theorem 5 is involved by Definition 4, we can restrict ourselves to the case that the sequence a_n is bounded.

Let us suppose first that a is an interelement of the bounded sequence a_n, and that a further element b is given. Then Theorem 2 involves that

$$u^* < ab \tag{1}$$

whenever u^* is a subelement of the sequence a_nb. If on the other side o^* is a superelement of the sequence a_nb, and o a superelement of the given sequence a_n, we have for $n \geqq k$, if k is an appropriate natural number,

$$o^* \vee (o + ob) > a_nb \vee (a_n + a_nb) = a_n,$$

hence, according to the supposed property of a,

$$o^* \vee (o + ob) > a$$

and

$$o^* > ab. \tag{2}$$

(1) and (2) involve that ab is an interelement of the sequence a_nb.

Now let us suppose conversely that ab is an interelement of the sequence a_nb for every element b. From this supposition it yields especially that $a(a \vee c)$ is an interelement of the sequence $a_n(a \vee c)$ whenever c is an upper bound of the sequence a_n. Thus we obtain immediately that a is an interelement of the sequence a_n as we want to prove.

THEOREM 6. *If a_n and a are elements of a Boolean ring A, then a is an interelement of the sequence a_n with respect to A if and only if $a \vee b$ is an interelement of the sequence $a_n \vee b$ with respect to A for every element b of A.*

PROOF. If $a \vee b$ is an interelement of the sequence $a_n \vee b$ for every element b then a is an interelement of the sequence a_n since we may put especially $b = 0$. Now let us suppose conversely that a is an interelement of the sequence a_n, and b an arbitrary element. Let us further suppose that the sequence a_n is *bounded*. In this case the sequence $a_n \vee b$ is bounded too. Let u^* be a subelement, and o^* a superelement of this sequence. Then $u^* + u^*b$ is a subelement of the sequence $a_n + a_nb$, thus also a subelement of the sequence a_n. According to the supposed property of the element a we have

$$u^* + u^*b < a,$$

whence

$$u^* < a \vee b. \tag{3}$$

On the other side Theorem 4 involves the inequality

$$o^* > a \vee b. \tag{4}$$

From (3) and (4) it follows that $a \vee b$ is an interelement of the sequence $a_n \vee b$.

If the sequence a_n is not bounded we observe that ac is, by Definition 4, an interelement of the bounded sequence a_nc for every element c. Hence, according to what we have just proved, $ac \vee bc = (a \vee b)c$ is an interelement of the sequence $a_nc \vee bc = (a_n \vee b)c$. Thus we obtain again the result that $a \vee b$ is an interelement of the sequence $a_n \vee b$.

THEOREM 7. *The interelements of a sequence of elements a_n of a Boolean ring*

A constitute a dense sublattice of A. (This sublattice is a Boole-Schröder lattice in the sense of the paper (IV).)

THEOREM 8. *If u is a subelement, and a an interelement of the same sequence a_n, then $u < a$.*

THEOREM 9. *If a is an interelement of a sequence a_n, and*

$$u < a^* < a$$

whenever u is a subelement of the same sequence, then a^* is as well an interelement of this sequence.

Theorems 7, 8, and 9 follow easily from Theorem 2 and from Definitions 3 and 4. (Notice that $a^{(1)}b < a^{(2)}b$ for every b implies $a^{(1)} < a^{(2)}$.)

DEFINITION 5. *If a_n is a sequence of elements of a Boolean ring A, and the lattice of the interelements of the sequence a_n with respect to A has a zero, this zero is called the lower limit of the sequence a_n with respect to A and denoted by* $^{(A)}\underline{\lim}_{n\to\infty} a_n$.

The expression "lower limit of the sequence a_n with respect to A" may be replaced by the simpler expression "lower limit of the sequence a_n" if there is no doubt to which Boolean ring it refers. In this case the sign $^{(A)}\underline{\lim}_{n\to\infty} a_n$ may be replaced by the simpler sign $\underline{\lim}_{n\to\infty} a_n$.

If $\underline{\lim}_{n\to\infty} a_n$ exists the lattice of the interelements of the sequence a_n is a Boolean ring, and $\underline{\lim}_{n\to\infty} a_n$ is its zero-element.

LEMMA 1. *If $a > u$ whenever u is a subelement of a given sequence a_n, and $\underline{\lim}_{n\to\infty} a_n$ exists, then*

$$a > \underline{\lim}_{n\to\infty} a_n .$$

PROOF. If u is a subelement of the sequence a_n then from the hypothesis and from Theorem 8 follows that

$$a \underline{\lim}_{n\to\infty} a_n > u.$$

This inequality implies, by Theorem 9, that $a \underline{\lim}_{n\to\infty} a_n$ is an interelement of the sequence a_n. Hence we have, according to Definition 5,

$$a \underline{\lim}_{n\to\infty} a_n > \underline{\lim}_{n\to\infty} a_n$$

or equivalently

$$a > \underline{\lim}_{n\to\infty} a_n$$

as we wished to prove.

DEFINITION 6. *If a_n is a sequence of elements of a Boolean ring A, and the lattice of the interelements of the sequence a_n with respect to A has a unit, this unit is called the upper limit of the sequence a_n with respect to A, and denoted by* $^{(A)}\overline{\lim}_{n\to\infty} a_n$.

The expression "upper limit of the sequence a_n with respect to A" may be replaced by the simpler expression "upper limit of the sequence a_n" if there is no doubt to which Boolean ring it refers. In this case the sign $\overline{\lim}^{(A)}_{n\to\infty} a_n$ may be replaced by the simpler sign $\overline{\lim}_{n\to\infty} a_n$.

LEMMA 2. *If $a < o$ whenever o is a superelement of a given bounded sequence a_n, and $\overline{\lim}_{n\to\infty} a_n$ exists, then*

$$a < \overline{\lim_{n\to\infty}} a_n .$$

PROOF. If o is a superelement of the sequence a_n then from the hypothesis follows that

$$a \vee \overline{\lim_{n\to\infty}} a_n < o.$$

Hence $a \vee \overline{\lim}_{n\to\infty} a_n$ is, by Definition 3, an interelement of the sequence a_n. On the other side we have, by Definition 6,

$$a \vee \overline{\lim_{n\to\infty}} a_n < \overline{\lim_{n\to\infty}} a_n$$

or equivalently

$$a < \overline{\lim_{n\to\infty}} a_n$$

as we wished to prove.

THEOREM 10. *If both $\underline{\lim}_{n\to\infty} a_n$ and $\overline{\lim}_{n\to\infty} a_n$ exist then*

$$\underline{\lim_{n\to\infty}} a_n < \overline{\lim_{n\to\infty}} a_n .$$

THEOREM 11. *If both $\underline{\lim}_{n\to\infty} a_n$ and $\overline{\lim}_{n\to\infty} a_n$ exist then a is an interelement of the sequence a_n if and only if*

$$\underline{\lim_{n\to\infty}} a_n < a < \overline{\lim_{n\to\infty}} a_n .$$

We need now some transfinite extensions of the known rules which the finite operations of a Boolean ring obey.

THEOREM 12. *Let $\mathfrak{B}$ be a class of elements of a Boolean ring A, and also b an element of A. If in this case $\prod_{a\in\mathfrak{B}} a$ exists then $\prod_{a\in\mathfrak{B}} (a \vee b)$ exists too and is equal to $(\prod_{a\in\mathfrak{B}} a) \vee b$.*

PROOF. It is obviously sufficient to prove that

$$c < a \vee b \tag{5}$$

for every element a of $\mathfrak{B}$ implies

$$c < (\prod_{a\in\mathfrak{B}} a) \vee b.$$

From (5) follows that

$$c + cb < a$$

for every element a of $\mathfrak{B}$. Hence

$$c + cb < \prod_{a \in \mathfrak{B}} a,$$

and

$$c < (\prod_{a \in \mathfrak{B}} a) \vee b$$

as we wished to prove.

In the following Theorems 13 till 17 the signs $\mathfrak{B}$ and b shall have the same significance as in Theorem 12. In the special case of Boolean σ-rings they were proved by J. von Neumann, "Lectures on continuous geometries," Princeton, 1937, Appendix, p. 7. See also H. M. Mac Neille, l.c., Theorem 7.19.

THEOREM 13. *If* $\sum_{a \in \mathfrak{B}} a$ *exists then* $\sum_{a \in \mathfrak{B}} ab$ *exists too and is equal to* $(\sum_{a \in \mathfrak{B}} a)b$.

PROOF. If $c > ab$ for every element a of $\mathfrak{B}$ then

$$c \vee [\sum_{x \in \mathfrak{B}} x + (\sum_{x \in \mathfrak{B}} x)b] > ab \vee [\sum_{x \in \mathfrak{B}} x + (\sum_{x \in \mathfrak{B}} x)b] > a \text{ for } a \in \mathfrak{B};$$

hence

$$c \vee [\sum_{x \in \mathfrak{B}} x + (\sum_{x \in \mathfrak{B}} x)b] > \sum_{a \in \mathfrak{B}} a,$$

and

$$c > (\sum_{a \in \mathfrak{B}} a)\, b.$$

THEOREM 14. *If* $\prod_{a \in \mathfrak{B}} a$ *exists then* $\sum_{a \in \mathfrak{B}} (b + ba)$ *exists too and is equal to* $b + b \prod_{a \in \mathfrak{B}} a$.

PROOF. It is clear that

$$b + b \prod_{a \in \mathfrak{B}} a > b + bx$$

for $x \in \mathfrak{B}$. If $c > b + ba$ for every element a of $\mathfrak{B}$ then

$$a > b + bc$$

for $a \in \mathfrak{B}$; hence

$$\prod_{a \in \mathfrak{B}} a > b + bc,$$

and

$$c > b + b \prod_{a \in \mathfrak{B}} a.$$

THEOREM 15. *If* $\sum_{a \in \mathfrak{B}} a$ *exists then* $\prod_{a \in \mathfrak{B}} (b + ba)$ *exists too and is equal to* $b + b \sum_{a \in \mathfrak{B}} a$.

PROOF. If $c < b + ba$ for $a \in \mathfrak{B}$ then $c < b$ and $ca = 0$ for $a \in \mathfrak{B}$; thus, by Theorem 13, $c \sum_{a \in \mathfrak{B}} a = 0$, and $c < b + b \sum_{a \in \mathfrak{B}} a$.

THEOREM 16. *If* $\prod_{a \in \mathfrak{B}} a$ *exists then* $\prod_{a \in \mathfrak{B}} (a + ab)$ *exists too and is equal to* $\prod_{a \in \mathfrak{B}} a + (\prod_{a \in \mathfrak{B}} a)b$.

PROOF. If $c < a + ab$ for $a \in \mathfrak{B}$ then $c < a$ for $a \in \mathfrak{B}$, thus $c < \prod_{a \in \mathfrak{B}} a$; on the other side $cb = 0$; hence

$$c < \prod_{a \in \mathfrak{B}} a + (\prod_{a \in \mathfrak{B}} a)\, b.$$

THEOREM 17. *If* $\sum_{a \in \mathfrak{B}} a$ *exists then* $\sum_{a \in \mathfrak{B}} (a + ab)$ *exists too and is equal to* $\sum_{a \in \mathfrak{B}} a + (\sum_{a \in \mathfrak{B}} a)b$.

PROOF. If $c > a + ab$ for $a \in \mathfrak{B}$ then $c \vee b > a$ for $a \in \mathfrak{B}$, thus $c \vee b > \sum_{a \in \mathfrak{B}} a$, and $c > \sum_{a \in \mathfrak{B}} a + (\sum_{a \in \mathfrak{B}} a)b$.

THEOREM 18. *If* $\mathfrak{a}$ *is an ideal in a Boolean ring* A, *then* $\mathfrak{a}''$ *is the class of all elements* b *of* A *such that* $b = \sum^{(A)}_{a \in \mathfrak{B}} a$ *for an appropriate subclass* $\mathfrak{B}$ *of* $\mathfrak{a}$. $\mathfrak{a}''$ *is a principal ideal if and only if* $\sum^{(A)}_{a \in \mathfrak{a}} a$ *exists; in this case* $\mathfrak{a}'' = \mathfrak{a}(\sum^{(A)}_{a \in \mathfrak{a}} a)$. ADDED IN PROOF: Theorem 18 is partly contained in Theorem 2.1 of M. H. Stone's paper on *Algebraic characterizations of special Boolean rings*, Fundamenta Mathematicae, vol. 29 (1937), pp. 223–303.

PROOF. From the definition of $\mathfrak{a}''$ and from Theorem 13 follows that $\mathfrak{a}''$ contains the element $\sum^{A}_{a \in \mathfrak{B}} a$ whenever $\mathfrak{B} \subset \mathfrak{a}$, and $\sum^{A}_{a \in \mathfrak{B}} a$ exists. If b is an arbitrary element of $\mathfrak{a}''$ we put $\mathfrak{B} = \mathfrak{a} \cdot \mathfrak{a}(b)$. Then $b > a$ whenever a is an element of $\mathfrak{B}$. If c is an arbitrary element of A then

$$(b + bc)x \in \mathfrak{B}$$

for all elements x of $\mathfrak{a}$. Hence if $c \in A$ and $c > a$ for every element a of $\mathfrak{B}$ we have, for $x \in \mathfrak{a}$,

$$(b + bc)x < c,$$

thus

$$(b + bc)x = 0$$

or

$$b + bc \in \mathfrak{a}'.$$

Since $b \in \mathfrak{a}''$ we have $b + bc = 0$, thus $b < c$. Hence

$$b = \sum_{a \in \mathfrak{B}}^{(A)} a.$$

This proves the first part of the theorem. The second part is now obvious.

THEOREM 19. $\varliminf_{n\to\infty} a_n$ *exists if and only if* $\sum_{u \in \mathfrak{U}} u$, $\mathfrak{U}$ *being the class of the subelements of the sequence* a_n, *exists; in this case*

$$\varliminf_{n\to\infty} a_n = \sum_{u \in \mathfrak{U}} u. \tag{6}$$

PROOF. If $\varliminf_{n\to\infty} a_n$ exists, from Lemma 1 and Theorem 8 follows immediately that $\sum_{u \in \mathfrak{U}} u$ exists, and the equation (6) holds. If $\sum_{u \in \mathfrak{U}} u$ exists then there exists also $\sum_{u \in \mathfrak{U}} ub$ for every element b. (See Theorem 13.) If we notice Theorem 2, we conclude easily that $\sum_{u \in \mathfrak{U}} ub$ is an interelement of the bounded sequence $a_n b$. On the other side we have

$$\sum_{u \in \mathfrak{U}} ub = \left(\sum_{u \in \mathfrak{U}} u\right) b.$$

Hence $\sum_{u \in \mathfrak{U}} u$ is an interelement of the sequence a_n. If a is an arbitrary interelement of this sequence, we have, by Theorem 8,

$$a > u$$

for every element u of $\mathfrak{U}$, and consequently

$$a > \sum_{u \in \mathfrak{U}} u.$$

Hence $\varliminf_{n\to\infty} a_n$ exists and is equal to $\sum_{u \in \mathfrak{U}} u$.

THEOREM 20. $\varliminf_{n\to\infty} a_n$ *exists if and only if* $\mathfrak{U}''$, $\mathfrak{U}$ *having the same significance as in Theorem* 19, *is a principal ideal* $\mathfrak{a}(a)$; *in this case*

$$\varliminf_{n\to\infty} a_n = a.$$

PROOF. See Theorems 18 and 19.

THEOREM 21. (Corollary of Theorem 20.) *If the ideal of the subelements of a sequence* a_n *is a principal ideal* $\mathfrak{a}(a)$, *then* $\varliminf_{n\to\infty} a_n$ *exists and is equal to* a.

THEOREM 22. *The upper limit of a bounded sequence* a_n *exists if and only if* $\prod_{o \in \mathfrak{O}} o$, $\mathfrak{O}$ *being the class of the superelements of the sequence* a_n, *exists; in this case*

$$\varlimsup_{n\to\infty} a_n = \prod_{o \in \mathfrak{O}} o.$$

The proof can be left to the reader. (Notice Lemma 2.)

THEOREM 23. (Corollary of Theorem 22.) *If the dual ideal of the superelements of a bounded sequence* a_n *has a zero then* $\varlimsup_{n\to\infty} a_n$ *exists and is equal to this zero.*

In this case the dual ideal of the superelements of the sequence a_n is a Boolean ring, and $\varlimsup_{n\to\infty} a_n$ is its zero.

LEMMA 3. *If* $\underline{\lim}_{n\to\infty} a_n$ *exists, and* b *is an arbitrary element, then* $\underline{\lim}_{n\to\infty} a_n b$ *exists too and is equal to* $b \underline{\lim}_{n\to\infty} a_n$.

PROOF. See Theorems 2, 13, and 19.

LEMMA 4. *If* $\overline{\lim}_{n\to\infty} a_n$ *exists, and* b *is an arbitrary element, then* $\overline{\lim}_{n\to\infty} a_n b$ *exists too and is equal to* $b \overline{\lim}_{n\to\infty} a_n$.

PROOF. Let $\overline{\lim}_{n\to\infty} a_n$ be denoted by a, and let c be an interelement of the sequence $a_n b$. Further let d be an arbitrary element, and o a superelement of the sequence $a_n d$. Thus o is also a superelement of the sequence $a_n bd$, and $cd < o$. We have also $ad < o$ because ad is an interelement of the sequence $a_n d$. Hence

$$(a \vee c)d < o.$$

If on the other side u is a subelement of the sequence $a_n d$, then $ad > u$, whence

$$(a \vee c)d > u.$$

We find that $a \vee c$ is an interelement of the sequence a_n. This involves, by Definition 6, $a \vee c < a$, or equivalently $c < a$. Besides we have obviously $c < b$. Hence $c < ab$ or

$$c < b \overline{\lim}_{n\to\infty} a_n.$$

Since $b \overline{\lim}_{n\to\infty} a_n$ is itself an interelement of the sequence $a_n b$, Lemma 4 is proved.

DEFINITION 7. *If a sequence of elements* a_n *of a Boolean ring* A *has exactly one interelement* a *with respect* A, *we say that the sequence* a_n *converges towards* a *with respect to* A, *or* a *is its limit with respect to* A, *and write*

$$^{(A)}\lim_{n\to\infty} a_n = a.$$

The sentences "The sequence a_n converges towards a with respect to A" and "a is the limit of the sequence a_n with respect to A" may be replaced respectively by the shorter sentences "The sequence a_n converges towards a" and "a is the limit of the sequence a_n" if there is no doubt to which Boolean ring they refer. In this case the sign $^{(A)}\lim_{n\to\infty} a_n$ may be replaced by the simpler sign $\lim_{n\to\infty} a_n$.

DEFINITION 8. *The sequential topology in a Boolean ring* A, *introduced by Definition* 7, *shall be denoted by* $\tau(A)$.

THEOREM 24. *A sequence of elements* a_n *of a Boolean ring is convergent if and only if* $\underline{\lim}_{n\to\infty} a_n$ *and* $\overline{\lim}_{n\to\infty} a_n$ *exist and are equal to one another. In this case*

$$\lim_{n\to\infty} a_n = \underline{\lim}_{n\to\infty} a_n = \overline{\lim}_{n\to\infty} a_n.$$

LEMMA 5. *If* $\lim_{n\to\infty} a_n$ *exists, and* b *is an arbitrary element, then* $\lim_{n\to\infty} a_n b$ *exists too and is equal to* $b \lim_{n\to\infty} a_n$.

PROOF. See Lemmas 3 and 4 and Theorem 24.

THEOREM 25. *If $a_n = a$ for every n then also*

$$\lim_{n\to\infty} a_n = a.$$

THEOREM 26. *A subelement (superelement) of a given sequence is also a subelement (superelement) of any partial sequence.*

THEOREM 27. *Any interelement of a partial sequence of a given sequence is also an interelement of the whole sequence.*

Theorem 26 is evident, Theorem 27 can be easily deduced from Theorem 26.

THEOREM 28. *If a_n is an arbitrary sequence of elements of a Boolean ring, and n_k $(k = 1, 2, 3, \cdots)$ are natural numbers different from each other, then there hold always those of the inclusion-relations*

$$\underline{\lim_{k\to\infty}}\, a_{n_k} > \underline{\lim_{n\to\infty}}\, a_n,$$

$$\underline{\lim_{k\to\infty}}\, a_{n_k} < \overline{\lim_{n\to\infty}}\, a_n,$$

$$\overline{\lim_{k\to\infty}}\, a_{n_k} > \underline{\lim_{n\to\infty}}\, a_n,$$

and

$$\overline{\lim_{k\to\infty}}\, a_{n_k} < \overline{\lim_{n\to\infty}}\, a_n,$$

in which both sides exist.

PROOF. See Definitions 5 and 6 and Theorem 27.

THEOREM 29. *If $\lim_{n\to\infty} a_n$ exists, and n_k $(k = 1, 2, 3, \cdots)$ are natural numbers different from each other, then $\lim_{k\to\infty} a_{n_k}$ exists too and is equal to $\lim_{n\to\infty} a_n$.*

PROOF. Let us suppose first that the sequence a_n is bounded. Then, if u is a subelement and o a superelement of this sequence, and u^* and o^* have the same meaning for the sequence a_{n_k} $(k = 1, 2, 3, \cdots)$, we have, by Theorem 26,

$$u^* < o$$

and

$$u < o^*.$$

Hence we have, by Lemmas 1 and 2,

$$u^* < \overline{\lim_{n\to\infty}}\, a_n$$

and

$$\underline{\lim_{n\to\infty}}\, a_n < o^*,$$

i.e.

$$u^* < \lim_{n\to\infty} a_n < o^*.$$

We see that $\lim_{n\to\infty} a_n$ is an interelement of the sequence a_{n_k} $(k = 1, 2, 3, \cdots)$.

The same result is obtained also in the case that the sequence a_n is not bounded. For if b is an arbitrary element Lemma 5 yields that $\lim_{n\to\infty} a_n b$ exists and is equal to $b \lim_{n\to\infty} a_n$; hence $b \lim_{n\to\infty} a_n$ is, by what we have just proved, an interelement of the sequence $a_{n_k} b$ $(k = 1, 2, 3, \cdots)$.

$\lim_{n\to\infty} a_n$ is also the only interelement of the sequence a_{n_k} $(k = 1, 2, 3, \cdots)$; for if the sequence a_{n_k} had an interelement a different from $\lim_{n\to\infty} a_n$ then a would be, by Theorem 27, also an interelement of the original sequence a_n , contrary to the hypothesis that this sequence is convergent. By this Theorem 29 is proved.

It can occur that both $\underline{\lim}_{n\to\infty} a_n$ and $\overline{\lim}_{n\to\infty} a_n$ exist but that there is neither a partial sequence of the sequence a_n converging towards $\underline{\lim}_{n\to\infty} a_n$ nor such a sequence converging towards $\overline{\lim}_{n\to\infty} a_n$. Let a and b be two elements such that neither $a < b$ nor $b < a$. Now for $n = 1, 2, 3, \cdots$ let $a_{2n-1} = a$, $a_{2n} = b$. Then we have obviously

$$\underline{\lim}_{n\to\infty} a_n = ab$$

and

$$\overline{\lim}_{n\to\infty} a_n = a \vee b.$$

Nevertheless there is neither a partial sequence of the given sequence converging towards ab nor such a sequence converging towards $a \vee b$.

THEOREM 30. *The union of a finite number of convergent sequences with the same limit is again a convergent sequence with the same limit.*

PROOF. Suppose that $a_n^{(1)}$ and $a_n^{(2)}$ are two convergent sequences, and that

$$\lim_{n\to\infty} a_n^{(1)} = \lim_{n\to\infty} a_n^{(2)} = a.$$

Suppose moreover that the sequence a_n consists of the terms of the sequences $a_n^{(1)}$ and $a_n^{(2)}$ in some succession. Then from Theorem 27 follows that a is an interelement of the sequence a_n . Now let a^* be an arbitrary interelement of this sequence, b an arbitrary element, $u^{(1)}$ and $u^{(2)}$ subelements, and $o^{(1)}$ and $o^{(2)}$ superelements of the sequences $a_n^{(1)}b$ and $a_n^{(2)}b$ respectively. Then $u^{(1)}u^{(2)}$ is a subelement, and $o^{(1)} \vee o^{(2)}$ a superelement of the sequence a_n . Since a^*b is an interelement of this sequence we have

$$u^{(1)}u^{(2)} < a^*b < o^{(1)} \vee o^{(2)}. \tag{7}$$

The sequences $a_n^{(1)}b$ and $a_n^{(2)}b$ are convergent, and

$$\lim_{n\to\infty} a_n^{(1)} b = \lim_{n\to\infty} a_n^{(2)} b = ab.$$

(See Lemma 5.) Hence we get from (7) by using twice Theorem 19 together with Theorem 13, and Theorem 22 together with Theorem 12

$$ab \cdot ab < a^*b < ab \vee ab$$

i.e.

$$a^*b = ab.$$

Since b is arbitrary we obtain the result $a^* = a$. This proves the theorem.

3. Connection with the Usual Convergence-Definition

The following results show that our definition specializes in the case of σ-lattices to the usual definition (cf. for example G. Birkhoff, II, p. 453).

THEOREM 31. *If $a_k > a_l$ whenever $k < l$, the sequence a_n is either convergent or has no interelement; it is convergent if and only if $\prod_{k=1}^{\infty} a_k$ exists; in this case we have*

$$\prod_{k=1}^{\infty} a_k = \lim_{n\to\infty} a_n . \tag{8}$$

PROOF. If a is an interelement of the sequence a_n, we have

$$a < a_n$$

for every n because all terms of the sequence are superelements of the sequence. If $u < a_n$ for every n then u is a subelement of the sequence, and we have

$$a > u.$$

Hence $\prod_{k=1}^{\infty} a_k$ exists and is equal to a. Thus we see that a is the only interelement of our sequence, is to be denoted by $\lim_{n\to\infty} a_n$, and the equation (8) holds.

If conversely $\prod_{k=1}^{\infty} a_k$ exists then it can be easily concluded that

$$u < \prod_{k=1}^{\infty} a_k < o$$

whenever u is a subelement and o a superelement of the sequence a_n, so that $\prod_{k=1}^{\infty} a_k$ is an interelement of this sequence.

THEOREM 32. (Corollary of Theorem 31.) *The sequence $\prod_{k=1}^{n} a_k$ ($n = 1, 2, 3, \cdots$) is either convergent or has no interelement; it is convergent if and only if $\prod_{k=1}^{\infty} a_k$ exists; in this case we have*

$$\prod_{k=1}^{\infty} a_k = \lim_{n\to\infty} \prod_{k=1}^{n} a_k .$$

THEOREM 33. *If $a_k < a_l$ whenever $k < l$, the sequence a_n is either convergent or has no interelement; it is convergent if and only if $\sum_{k=1}^{\infty} a_k$ exists; in this case we have*

$$\sum_{k=1}^{\infty} a_k = \lim_{n\to\infty} a_n .$$

If the suppositions of Theorem 33 are satisfied, and the sequence a_n has an interelement a, then we have

$$a > a_n$$

for every n because all terms of the sequence are subelements of the sequence. Hence the sequence a_n is certainly bounded. The same is also true if $\sum_{k=1}^{\infty} a_k$ exists. After this is settled Theorem 33 can be proved in a similar way as Theorem 31.

THEOREM 34. (Corollary of Theorem 33.) *The sequence* $\sum_{k=1}^{n} a_k$ $(n = 1, 2, 3, \cdots)$ *is either convergent or has no interelement; it is convergent if and only if* $\sum_{k=1}^{\infty} a_k$ *exists; in this case*

$$\sum_{k=1}^{\infty} a_k = \lim_{n\to\infty} \sum_{k=1}^{n} a_k .$$

THEOREM 35. *Any monotonic sequence of elements of a Boolean σ-ring is convergent.*

Theorem 35 follows immediately from Theorems 31 and 33.

THEOREM 36. *If for* $n = 1, 2, 3, \cdots$ $\prod_{k=n}^{\infty} a_k$ *exists then* $\underline{\lim}_{n\to\infty} a_n$ *exists if and only if* $\sum_{n=1}^{\infty} \prod_{k=n}^{\infty} a_k$ *exists; in this case we have*

$$\underline{\lim_{n\to\infty}}\, a_n = \sum_{n=1}^{\infty} \prod_{k=n}^{\infty} a_k = \lim_{n\to\infty} \prod_{k=n}^{\infty} a_k .$$

PROOF. Definition 1 involves that to every subelement u of the sequence a_n there can be stated a natural number m so that

$$\prod_{k=m}^{\infty} a_k > u.$$

From this it can be easily concluded that $\sum_{u \in \mathfrak{U}} u$ exists if and only if $\sum_{n=1}^{\infty} \prod_{k=n}^{\infty} a_k$ exists, and that in this case both elements are identical. ($\mathfrak{U}$ is the ideal of the subelements of the sequence a_n.) Now Theorems 19 and 33 show that Theorem 36 is true.

THEOREM 37. *If for* $n = 1, 2, 3, \cdots$ $\sum_{k=n}^{\infty} a_k$ *exists then* $\overline{\lim}_{n\to\infty} a_n$ *exists if and only if* $\prod_{n=1}^{\infty} \sum_{k=n}^{\infty} a_k$ *exists; in this case we have*

$$\overline{\lim_{n\to\infty}}\, a_n = \prod_{n=1}^{\infty} \sum_{k=n}^{\infty} a_k = \lim_{n\to\infty} \sum_{k=n}^{\infty} a_k .$$

Theorem 37 can be proved by the aid of Theorems 22 and 31 in a similar way as Theorem 36 by the aid of Theorems 19 and 33.

THEOREM 38. *If* a_n *is a sequence of elements of a Boolean σ-ring then there exist certainly both* $\underline{\lim}_{n\to\infty} a_n$ *and* $\overline{\lim}_{n\to\infty} a_n$.

PROOF. See Theorems 36 and 37.

THEOREM 39. *If a_n is a sequence of elements of a Boolean σ-ring then* $\lim_{n\to\infty} a_n$ *exists if and only if*

$$\sum_{n=1}^{\infty}\prod_{k=n}^{\infty} a_k = \prod_{n=1}^{\infty}\sum_{k=n}^{\infty} a_k .$$

In this case we have

$$\lim_{n\to\infty} a_n = \sum_{n=1}^{\infty}\prod_{k=n}^{\infty} a_k = \prod_{n=1}^{\infty}\sum_{k=n}^{\infty} a_k .$$

Thus our Definition 7 coincides in the case of a Boolean σ-ring with the definition of convergence and limit given in the introduction.

LEMMA 6. *If a_n is a sequence of elements of a Boolean σ-ring with unit then*

$$\underline{\lim_{n\to\infty}}\, a_n' = (\overline{\lim_{n\to\infty}}\, a_n)',$$

$$\overline{\lim_{n\to\infty}}\, a_n' = (\underline{\lim_{n\to\infty}}\, a_n)';$$

if moreover the sequence a_n is convergent, the sequence a_n' is convergent too, and

$$\lim_{n\to\infty} a_n' = (\lim_{n\to\infty} a_n)'.$$

Lemma 6 can be easily deduced from Theorems 14, 15, 36, and 37.

LEMMA 7. *If a_n and a are elements of a Boolean σ-ring with unit, and a is an interelement of the sequence a_n, then a' is an interelement of the sequence a_n'.*

PROOF. See Theorem 11 and Lemma 6.

THEOREM 40. *If $a_n > c$ for every n then the equation*

$$\underline{\lim_{n\to\infty}}\, a_n = c \tag{9}$$

holds if and only if for every n

$$\prod_{k=n}^{\infty} a_k = c. \tag{10}$$

PROOF. If (9) holds, and the element u is of the property that for $k \geqq n$ $u < a_k$, then $u < c$ because u is a subelement of the sequence a_n. This proves (10). If (10) holds then (9) follows immediately from Theorem 36.

THEOREM 41. *If $a_n < c$ for every n then the equation*

$$\overline{\lim_{n\to\infty}}\, a_n = c$$

holds if and only if for every n

$$\sum_{k=n}^{\infty} a_k = c.$$

THEOREM 42. *An element a is an interelement of a sequence a_n if and only if for every n*

$$\sum_{k=n}^{\infty} a_k a = a, \tag{11}$$

and

$$\prod_{k=n}^{\infty} (a_k \vee a) = a. \tag{12}$$

PROOF. If a is an interelement of the sequence a_n then it is, by Theorems 5 and 6, also an interelement of the sequences $a_n a$ and $a_n \vee a$. On the other side it is a superelement of the sequence $a_n a$ and a subelement of the sequence $a_n \vee a$. Hence we have

$$\overline{\lim_{n\to\infty}}\, a_n a = a,$$

$$\underline{\lim_{n\to\infty}}\, (a_n \vee a) = a.$$

Now Theorems 41 and 40 involve that the equations (11) and (12) hold.

Let us suppose conversely that the equations (11) and (12) hold. Let b be an arbitrary element, and u a subelement and o a superelement of the sequence $a_n b$. Further let m be a natural number of the property that

$$u < a_k b < o$$

for $k \geqq m$. Then we have

$$ab = \sum_{k=m}^{\infty} a_k ab < ao,$$

whence

$$ab < o.$$

Similarly we have

$$a = \prod_{k=m}^{\infty} (a_k \vee a) > u \vee a,$$

whence

$$u < a.$$

Since also $u < b$ we get

$$u < ab.$$

Thus we see that ab is an interelement of the sequence $a_n b$; hence a is an interelement of the sequence a_n as we wished to prove.

Another form of Theorem 42 is

THEOREM 43. *An element a is an interelement of a sequence a_n if and only if*

$$\overline{\lim_{n\to\infty}}\, a_n a = a, \tag{13}$$

and

$$\underline{\lim_{n\to\infty}}\, (a_n \vee a) = a. \tag{14}$$

PROOF. See Theorems 40, 41 and 42.

The following Theorems 44 and 45 will be needed in the next paragraph.

THEOREM 44. *The equation*

$$\underline{\lim_{n\to\infty}}\, a_n = a \tag{15}$$

is equivalent with the couple of the equations

$$\lim_{n\to\infty} a_n a = a \tag{16}$$

and

$$\underline{\lim_{n\to\infty}}\, (a_n \vee a) = a. \tag{14}$$

PROOF. That (15) implies (14) and (16) follows from Lemma 3 and Theorem 43. If (14) and (16) hold then a is, by Theorem 43, an interelement of the sequence a_n; if b is another such interelement then, by Theorem 5 and (16), $ab = a$, i.e. $a < b$.

THEOREM 45. *The equation*

$$\overline{\lim_{n\to\infty}}\, a_n = a \tag{17}$$

is equivalent with the couple of the equations

$$\overline{\lim_{n\to\infty}}\, a_n a = a \tag{13}$$

and

$$\lim_{n\to\infty} (a_n \vee a) = a. \tag{18}$$

PROOF. If (17) holds we have to prove, in face of Theorem 43, the equation (18) only. Let b be an arbitrary element, and o a superelement of the sequence $a_n b$. Then we have $o > ab$. If moreover c is an interelement of the sequence $a_n \vee a$ then bc is an interelement of the sequence $a_n b \vee ab$; since o is a superelement also of this sequence we have $bc < o$; thus, by Lemmas 2 and 4, $bc < ab$. Since b is arbitrary we have $c < a$. On the other side we have certainly $c > a$. Hence $c = a$. This proves (18).

If (13) and (18) hold then a is, by Theorem 43, an interelement of the sequence a_n; if b is another such interelement then, by Theorem 6 and (18), $a \vee b = a$, thus $a > b$. This proves (17).

4. Invariant subrings

If $\mathfrak{a}$ is a subring of a Boolean ring A, the restriction of the sequential topology $\tau(A)$ (Definition 8) to the subring $\mathfrak{a}$ is not necessarily identical with $\tau(\mathfrak{a})$.

Let for instance A be the Boolean algebra of all classes of natural numbers, a_n the class of all natural numbers divisible by n ($n = 1, 2, 3, \cdots$), and $\mathfrak{a}$ the subring of A generated by the elements a_n ($n = 1, 2, 3, \cdots$) of A. $\mathfrak{a}$ contains the unit of A. If a is a polynomial in the elements a_k ($k = 1, 2, \cdots, m$), and M the least common multiple of the natural numbers $1, 2, 3, \cdots, m$, then a natural number l is an element of a if and only if $l + M$ is an element of a. This yields that every element of $\mathfrak{a}$ is either an infinite class or the void class.

Now the classes $\prod_{k=n}^{\infty (A)} a_k' = \left(\sum_{k=n}^{\infty (A)} a_k\right)'$ are finite. Hence we have

$$\prod_{k=n}^{\infty (\mathfrak{a})} a_k' = 0$$

and

$$\sum_{k=n}^{\infty (\mathfrak{a})} a_k = e$$

if e is the unit of A. From the last equation follows

$$^{(\mathfrak{a})}\overline{\lim_{n\to\infty}}\, a_n = e.$$

It can also easily be proved that

$$^{(\mathfrak{a})}\underline{\lim_{n\to\infty}}\, a_n = 0.$$

We see that the sequence a_n is not convergent with respect to $\mathfrak{a}$. On the other hand we have obviously

$$^{(A)}\lim_{n\to\infty} a_n = 0.$$

Thus $\tau(\mathfrak{a})$ is not identical with the restriction of $\tau(A)$ to the subring $\mathfrak{a}$.

THEOREM 46. *If $\mathfrak{a}$ is a subring of a Boolean ring A, a_n ($n = 1, 2, 3, \cdots$) and a are elements of $\mathfrak{a}$, and a is an interelement of the sequence a_n with respect to A, then a is also an interelement of the sequence a_n with respect to $\mathfrak{a}$.*

PROOF. If b is an element of $\mathfrak{a}$ then ab is, by Theorem 5, an interelement of the sequence a_nb with respect to A. From Definition 3 it can be easily concluded that ab is also an interelement of the sequence a_nb with respect to $\mathfrak{a}$. Since this holds for every element b of $\mathfrak{a}$ Theorem 46 is proved.

THEOREM 47. *If $\mathfrak{a}$ is a subring of a Boolean ring A, $\mathfrak{B}$ a subclass of $\mathfrak{a}$, and b an element of $\mathfrak{a}$, then*

$$\prod_{a\in\mathfrak{B}}^{(A)} a = b$$

implies

$$\prod_{a \in \mathfrak{B}}^{(\mathfrak{a})} = b,$$

and

$$\sum_{a \in \mathfrak{B}}^{(A)} a = b$$

implies

$$\sum_{a \in \mathfrak{B}}^{(\mathfrak{a})} a = b.$$

DEFINITION 9. *A subring* $\mathfrak{a}$ *of a Boolean ring* A *is said to be invariant in* A *(*A *an invariant extension of* $\mathfrak{a}$, A *invariant over* $\mathfrak{a}$*) if the following condition is satisfied: if* $\mathfrak{B}$ *is a subclass of* $\mathfrak{a}$, *and* b *an element of* $\mathfrak{a}$, *then*

$$\sum_{a \in \mathfrak{B}}^{(\mathfrak{a})} a = b$$

implies

$$\sum_{a \in \mathfrak{B}}^{(A)} a = b.$$

Let for instance A be the Boolean algebra of all sets of real numbers and $\mathfrak{a}$ the subclass of A defined in the following way: a is an element of $\mathfrak{a}$ if and only if it is either the empty set or the union of a finite number of sets, every one of which is either an open real interval or consists of exactly one real number. Then it is easily to be seen that $\mathfrak{a}$ is an invariant subring of A.

Let, on the other hand, $\mathfrak{b}$ be the subclass of A defined in the following way: b is an element of $\mathfrak{b}$ if and only if it is either the empty set or the union of a finite number of right-hand open (and left-hand closed) real intervals. Then $\mathfrak{b}$ is as well a subring of A, but is not invariant in A. For if $c(\mu)$ (μ a positive real number) is the set of those real numbers ν which satisfy the inequality $0 \leqq \nu < \mu$ then $\prod_{\mu>0}^{(A)} c(\mu)$ is the set consisting only of the number zero while $\prod_{\mu>0}^{(\mathfrak{b})} c(\mu)$ is the empty set. (See the following Theorem 48.)

THEOREM 48. *If* $\mathfrak{a}$ *is an invariant subring of a Boolean ring* A, $\mathfrak{B}$ *a subclass of* $\mathfrak{a}$, *and* b *an element of* $\mathfrak{a}$, *then*

$$\prod_{a \in \mathfrak{B}}^{(\mathfrak{a})} a = b$$

implies

$$\prod_{a \in \mathfrak{B}}^{(A)} a = b.$$

PROOF. If a_0 is a fixed element of $\mathfrak{B}$, we have, by Theorem 14,

$$\sum_{a \in \mathfrak{B}}^{(\mathfrak{a})} (a_0 + a_0 a) = a_0 + a_0 \prod_{a \in \mathfrak{B}}^{(\mathfrak{a})} a = a_0 + b;$$

hence, by Definition 9,

$$\sum_{a \in \mathfrak{B}}^{(A)} (a_0 + a_0 a) = a_0 + b$$

and, by Theorem 15,

$$\prod_{a \in \mathfrak{B}}^{(A)} [a_0 + a_0(a_0 + a_0 a)] = a_0 + a_0(a_0 + b)$$

i.e.

$$\prod_{a \in \mathfrak{B}}^{(A)} a = b$$

as we wished to prove.

THEOREM 49. *If $\mathfrak{a}$ is an invariant subring of a Boolean ring A, a_n ($n = 1, 2, 3, \cdots$) and a are elements of $\mathfrak{a}$, and a is an interelement of the sequence a_n with respect to $\mathfrak{a}$, then a is also an interelement of the sequence a_n with respect to A.*

PROOF. See Theorem 42, Definition 9, and Theorem 48.

THEOREM 50. *If $\mathfrak{a}$ is an invariant subring of a Boolean ring A, and a_n ($n = 1, 2, 3, \cdots$) and a are elements of $\mathfrak{a}$, then the equation*

$$(19) \qquad {}^{(\mathfrak{a})}\lim_{n \to \infty} a_n = a$$

holds if and only if the equation

$$(20) \qquad {}^{(A)}\lim_{n \to \infty} a_n = a$$

holds.

In other words: if $\mathfrak{a}$ is an invariant subring of a Boolean ring A, then the sequential topology $\tau(\mathfrak{a})$ is the restriction of the sequential topology $\tau(A)$ to the subring $\mathfrak{a}$.

PROOF. From Definition 7 and from Theorems 46 and 49 follows immediately that (20) implies (19). Let us suppose now that (19) is satisfied. Let further b be an interelement of the sequence a_n with respect to A, and $\mathfrak{U}$ the class of those subelements of this sequence which are contained in $\mathfrak{a}$. Then, by Theorem 19,

$$\sum_{u \in \mathfrak{U}}^{(\mathfrak{a})} u = a;$$

hence, we have, by Definition 9, also

$$\sum_{u \in \mathfrak{U}}^{(A)} u = a$$

and, by Theorem 8,

$$b > a.$$

If the sequence a_n is bounded with respect to $\mathfrak{a}$, we can prove in a similar way that also $b < a$; thus we know that (19) implies (20) if the sequence a_n is bounded

with respect to $\mathfrak{a}$. If this condition is not satisfied, let c be an arbitrary element of $\mathfrak{a}$; we have, by Lemma 5,

$$^{(\mathfrak{a})}\lim_{n\to\infty} a_n c = ac$$

and, by what we have just proved,

$$^{(A)}\lim_{n\to\infty} a_n c = ac.$$

Hence we have

$$bc = ac$$

and

$$(b + ab)c = 0$$

for every element c of $\mathfrak{a}$. Thus we have especially

$$(b + ab)a_n = 0$$

for $n = 1, 2, 3, \cdots$. The last equation involves, by Theorem 5, the equation

$$(b + ab)b = 0$$

and consequently

$$b < a.$$

This proves the theorem also for the case that the sequence a_n is not bounded with respect to $\mathfrak{a}$.

From Theorem 50 follows that in the example given at the beginning of this paragraph the subring $\mathfrak{a}$ is not invariant in A.

THEOREM 51. *If $\mathfrak{a}$ is an invariant subring of a Boolean ring A, and a_n $(n = 1, 2, 3, \cdots)$ and a are elements of $\mathfrak{a}$, then the equation*

$$^{(\mathfrak{a})}\underline{\lim}_{n\to\infty} a_n$$

holds if and only if the equation

$$^{(A)}\underline{\lim}_{n\to\infty} a_n$$

holds.

PROOF. In face of Theorems 44 and 50 we have only to prove that the equation

$$\underline{\lim}_{n\to\infty} (a_n \vee a) = a \tag{14}$$

holds with respect to $\mathfrak{a}$ if and only if it holds with respect to A. But this assertion is equivalent with the assertion that a is an interelement of the sequence

$a_n \vee a$ with respect to $\mathfrak{a}$ if and only if it is with respect to A; and that the latter assertion is true follows from Theorems 46 and 49.

THEOREM 52. *If $\mathfrak{a}$ is an invariant subring of a Boolean ring A, and a_n $(n = 1, 2, 3, \cdots)$ and a are elements of $\mathfrak{a}$, then the equation*

$$^{(\mathfrak{a})}\overline{\lim_{n\to\infty}}\, a_n = a$$

holds if and only if the equation

$$^{(A)}\overline{\lim_{n\to\infty}}\, a_n = a$$

holds.

PROOF. Compare the proof of Theorem 51 and see Theorems 45, 50, 46, and 49.

THEOREM 53. *If $\mathfrak{a}_2$ is an invariant subring of a Boolean ring A, and $\mathfrak{a}_1$ an invariant subring of $\mathfrak{a}_2$, then $\mathfrak{a}_1$ is also an invariant subring of A.*

THEOREM 54. *Let $\mathfrak{B}$ be a chain of subrings of a Boolean ring A such that $\mathfrak{a} \in \mathfrak{B}$, $\mathfrak{b} \in \mathfrak{B}$, and $\mathfrak{a} \subset \mathfrak{b}$ imply that $\mathfrak{a}$ is invariant in $\mathfrak{b}$. Then each element $\mathfrak{a}$ of $\mathfrak{B}$ is also invariant in $\sum^{(\mathfrak{R})}_{\mathfrak{b}\in\mathfrak{B}} \mathfrak{b}$.* (That $\mathfrak{B}$ is a chain means that $\mathfrak{a} \in \mathfrak{B}$, $\mathfrak{b} \in \mathfrak{B}$ imply $\mathfrak{a} \subset \mathfrak{b}$ or $\mathfrak{b} \subset \mathfrak{a}$. $\mathfrak{R}$ is the lattice of all subrings of A.)

PROOF. Let $\mathfrak{a}$ be an element of $\mathfrak{B}$, $\mathfrak{C}$ a subset of $\mathfrak{a}$, and b an element of $\mathfrak{a}$. Suppose that

$$\sum_{a\in\mathfrak{C}}^{(\mathfrak{a})} a = b.$$

If c is an element of $\sum^{(\mathfrak{R})}_{\mathfrak{b}\in\mathfrak{B}} \mathfrak{b}$, and $c > a$ whenever $a \in \mathfrak{C}$, then c is contained in some element $\mathfrak{b}$ of $\mathfrak{B}$, and it can be supposed that $\mathfrak{b} \supset \mathfrak{a}$ so that $\mathfrak{a}$ is invariant in $\mathfrak{b}$. (Notice that $\sum^{(\mathfrak{R})}_{\mathfrak{b}\in\mathfrak{B}} \mathfrak{b}$ is the *set-theoretical* union of the subrings $\mathfrak{b}$ which are elements of $\mathfrak{B}$.) Hence $c > b$. This proves the theorem.

THEOREM 55. *If $\mathfrak{a}_1$ is an invariant subring of a Boolean ring A and $\mathfrak{a}_2$ a subring of A containing $\mathfrak{a}_1$, then $\mathfrak{a}_1$ is also invariant in $\mathfrak{a}_2$.*

PROOF. If $\mathfrak{B}$ is a subclass of $\mathfrak{a}_1$, b an element of $\mathfrak{a}_1$, and

$$\sum_{a\in\mathfrak{B}}^{(\mathfrak{a}_1)} a = b,$$

we have

$$\sum_{a\in\mathfrak{B}}^{(A)} a = b$$

because $\mathfrak{a}_1$ is invariant in A, and

$$\sum_{a\in\mathfrak{B}}^{(\mathfrak{a}_2)} a = b$$

by Theorem 47.

THEOREM 56. *Every ideal $\mathfrak{a}$ of a Boolean ring A is invariant in A.*

PROOF. Let $\mathfrak{B}$ be a subclass of $\mathfrak{a}$, b an element of $\mathfrak{a}$, and

$$\sum_{a \in \mathfrak{B}}^{(\mathfrak{a})} a = b.$$

If c is an arbitrary element of A, and $c > a$ whenever $a \in \mathfrak{B}$, we have also

$$bc > a$$

for every element a of $\mathfrak{B}$. Since bc is an element of $\mathfrak{a}$, we have

$$bc > b,$$

or equivalently

$$c > b$$

proving the theorem.

THEOREM 57. *If $\mathfrak{a}$ is an invariant subring of a Boolean ring A, and $\mathfrak{b}$ a normal ideal in A, then $\mathfrak{ab}$ is a normal ideal in $\mathfrak{a}$, and every normal ideal in $\mathfrak{a}$ can be obtained in this way.*

PROOF. If $\mathfrak{b}$ is a normal ideal in A, it is obvious that $\mathfrak{ab}$ is an ideal in $\mathfrak{a}$. If a is an element of the second orthocomplement of $\mathfrak{ab}$ with respect to $\mathfrak{a}$, we have, by Theorem 18,

$$a = \sum_{x \in \mathfrak{B}}^{(\mathfrak{a})} x$$

for an appropriate subclass $\mathfrak{B}$ of $\mathfrak{ab}$. Since $\mathfrak{a}$ is invariant in A, we have also

$$a = \sum_{x \in \mathfrak{B}}^{(A)} x.$$

Hence a is, again by Theorem 18, an element of $\mathfrak{b}$; since it is also an element of $\mathfrak{a}$, it is an element of $\mathfrak{ab}$. This proves the first part of the theorem.

If on the other side $\mathfrak{c}$ is an arbitrary normal ideal in $\mathfrak{a}$, we have obviously

$$\mathfrak{c} = \mathfrak{ac}''$$

the dashing referred to A, and $\mathfrak{c}''$ is a normal ideal in A.

5. Join-extensions

The relation between a partially ordered set and its completion by cuts (cf. H. MacNeille, "Partially ordered sets," Trans. Am. Math. Soc. 42 (1937), p. 445, line 9 from bottom) is a special case of the following definition.

DEFINITION 10. *If A is a subring of a Boolean ring R, then R is said to be a join-extension of A if to every element b of R there exists a subclass $\mathfrak{B}$ of A such that*

$$b = \sum_{a \in \mathfrak{B}}^{(R)} a.$$

If R is a join-extension of a Boolean ring A and a subring of a Boolean ring B, we shall say that R is a join-extension of A in B.

THEOREM 58. *If R is a join-extension of a Boolean ring A, then every element of R is the join of those elements of A which it contains, with respect to R.*

THEOREM 59. *If A and A^* are Boolean rings between which there exists an isomorphism Λ, R a join-extension of A, and R^* a join-extension of A^*, then there exists at the utmost one extension of the isomorphism Λ to an isomorphism Λ_R between R and R^*. If Λ_R exists it can be defined in the following way: let b be an element of R, $\mathfrak{a}$ the class of those elements of A which are contained in b, and $\mathfrak{a}^*$ the subclass of A^* corresponding to $\mathfrak{a}$ by virtue of Λ; then the element b^* of R^* corresponding to b by virtue of Λ_R is defined by the equation*

$$b^* = \sum_{a^* \in \mathfrak{a}^*}^{(R^*)} a^*.$$

The proof of Theorems 58 and 59 can be left to the reader.

THEOREM 60. (Corollary of Theorem 59.) *If R is a join-extension of a Boolean ring A, the identity is the only automorphism of R which carries every element of A into itself.*

THEOREM 61. *Any join-extension of a Boolean ring A is an invariant extension of A.*

PROOF. Let R be a join-extension of A, b an element of A, and $\mathfrak{B}$ a subclass of A such that

$$\sum_{a \in \mathfrak{B}}^{(A)} a = b. \tag{21}$$

Further let c be an element of R such that

$$c > a$$

for $a \in \mathfrak{B}$. Then we have also

$$(b + bc)a = 0 \tag{22}$$

for $a \in \mathfrak{B}$. On the other side $b + bc$ is itself an element of R. Hence we have, by Theorem 58,

$$b + bc = \sum_{\substack{x < b+bc \\ x \in A}}^{(R)} x. \tag{23}$$

From (22) and (23) follows

$$xa = 0 \tag{24}$$

for $x < b + bc$, $x \in A$, $a \in \mathfrak{B}$, and from (21) and (24), by Theorem 13,

$$xb = 0$$

for $x < b + bc$, $x \in A$. Now we get from (23)

$$b + bc = 0,$$

thus $c > b$. This proves

$$\sum_{a \in \mathfrak{B}}^{(R)} a = b.$$

THEOREM 62. *If R_1 is a join-extension of A, and R_2 a join-extension of R_1, then R_2 is a join-extension of A.*

PROOF. If b is an element of R_2 we have, by hypothesis and by Theorem 58,

$$b = \sum_{\substack{a<b \\ a \in R_1}}^{(R_2)} a$$

and

$$a = \sum_{\substack{x<a \\ x \in A}}^{(R_1)} x$$

for $a < b$, $a \in R_1$. Since R_1 is, by Theorem 61, invariant in R_2 we have also

$$a = \sum_{\substack{x<a \\ x \in A}}^{(R_2)} x \quad \text{for} \quad a < b, a \in R_1.$$

Hence

$$b = \sum_{\substack{a<b \\ a \in R_1}}^{(R_2)} \sum_{\substack{x<a \\ x \in A}}^{(R_2)} x.$$

This proves the theorem, according to Definition 10.

THEOREM 63. (Compare Theorem 54.) *Let $\mathfrak{B}$ be a chain of subrings of a Boolean ring B such that $R_1 \in \mathfrak{B}$, $R_2 \in \mathfrak{B}$ and $R_1 \subset R_2$ imply that R_2 is a join-extension of R_1. Then $\sum_{R \in \mathfrak{B}}^{(\mathfrak{R})} R$ is a join-extension of each element R of $\mathfrak{B}$, by $\mathfrak{R}$ denoted the lattice of all subrings of B.*

PROOF. If R_0 is an arbitrary element of $\mathfrak{B}$, and b an element of $\sum_{R \in \mathfrak{B}}^{(\mathfrak{R})} R$, then b is also contained in at least one subring R of B which is an element of $\mathfrak{B}$, and it can be supposed that $R_0 \subset R$. Hence we have

$$b = \sum_{\substack{a<b \\ a \in R_0}}^{(R)} a.$$

From this equation follows, by Theorem 54,

$$b = \sum_{\substack{a<b \\ a \in R_0}}^{(L)} a$$

if $\sum_{R \in \mathfrak{B}}^{(\mathfrak{R})} R$ is denoted by L. Hence $\sum_{R \in \mathfrak{B}}^{(\mathfrak{R})} R$ is a join-extension of R_0, proving the theorem.

THEOREM 64. *If R_2 is a join-extension of a Boolean ring A, and R_1 a subring of R_2 containing A, then R_1 is a join-extension of A, and R_2 is a join-extension of R_1.*

PROOF. If b is an element of R_1, it is also an element of R_2. Hence we have, by Theorem 58 and by hypothesis,

$$b = \sum_{\substack{a<b \\ a \in A}}^{(R_2)} a$$

and, by Theorem 47,

$$b = \sum_{\substack{a<b \\ a \in A}}^{(R_1)} a.$$

This proves that R_1 is a join-extension of A.

The second part of Theorem 64 is trivial.

If A and R are subrings of a Boolean ring B, A is invariant in B, and R a join-extension of A, then R is not necessarily invariant in B. Let B be the Boolean ring of all classes of natural numbers, A the subring of B consisting of all finite classes of natural numbers different from 1, and R the subring of B consisting of the elements of A and their complements with respect to the unit-element of B. Then A is invariant in B (it is an ideal of B), R is a join-extension of A, but R is not invariant in B.

THEOREM 65. *If A, R_1, and R_2 are subrings of a Boolean ring R_3, $A \subset R_1 \subset R_2$, and R_3 is a join-extension of A, then R_2 is a join-extension of R_1.*

PROOF. See Theorem 64.

THEOREM 66. *If A is an invariant subring of a Boolean ring B, and N the class of all elements b of B such that*

$$b = \sum_{x \in \mathfrak{B}}^{(B)} x$$

for an appropriate subclass $\mathfrak{B}$ of A (so that especially all elements of A belong to N) then N is itself a subring of B; hence N is a join-extension of A in B.

PROOF. It is obvious that $b_1 \in N$ and $b_2 \in N$ imply $b_1 b_2 \in N$ and $b_1 \vee b_2 \in N$. (See Theorem 13.)

Next we shall prove that $a \in A$ and $b \in N$ imply $a + ab \in N$. We put

$$\mathfrak{a}(a)\mathfrak{a}(b)A = \mathfrak{a},$$

the signs $\mathfrak{a}(a)$ and $\mathfrak{a}(b)$ referred to the Boolean ring B. $\mathfrak{a}(a)A$ is an ideal in A and consequently, by Theorem 56, invariant in A. Since A is, by hypothesis, invariant in B, $\mathfrak{a}(a)A$ is, by Theorem 53, as well invariant in B. Hence $\mathfrak{a}(a)\mathfrak{a}(b)A$ is, by Theorem 57, a normal ideal in $\mathfrak{a}(a)A$. Now we have

$$(\mathfrak{a} \vee_{\mathfrak{Z}} \mathfrak{a}')'' = \mathfrak{a}(a)A$$

if the dashes are referred to $\mathfrak{a}(a)A$, and $\mathfrak{Z}$ is the distributive lattice of the ideals in $\mathfrak{a}(a)A$. Consequently we have, by Theorem 18,

$$\sum_{x \in \mathfrak{a} \vee_{\mathfrak{Z}} \mathfrak{a}'}^{(\mathfrak{a}(a)A)} x = a$$

and

$$\sum_{\substack{x \in \mathfrak{a} \\ \text{or } x \in \mathfrak{a}'}}^{(\mathfrak{a}(a)A)} x = a$$

because every element of $\mathfrak{a} \vee_{\mathfrak{Z}} \mathfrak{a}'$ can be decomposed into the join of an element of $\mathfrak{a}$ and an element of $\mathfrak{a}'$. Hence we have also

$$\sum_{\substack{x \in \mathfrak{a} \\ \text{or } x \in \mathfrak{a}'}}^{(B)} x = a. \tag{25}$$

Since ab is an element of N, we have

$$ab = \sum_{x \in \mathfrak{B}}^{(B)} x$$

for an appropriate subclass $\mathfrak{B}$ of A. From this it can be easily concluded that also

$$ab = \sum_{\substack{x \in A \\ x<ab}}^{(B)} x,$$

i.e.

$$ab = \sum_{x \in \mathfrak{a}}^{(B)} x.$$

If on the other side $y \in \mathfrak{a}'$, then $y < a$ and $yx = 0$ for $x \in \mathfrak{a}$, whence $yb = 0$, and

$$y < a + ab.$$

If d is an element of B, and $d > y$ whenever $y \in \mathfrak{a}'$ then

$$d \vee ab > x$$

for $x \in \mathfrak{a}$ as well as for $x \in \mathfrak{a}'$; hence, according to (25),

$$d \vee ab > a,$$

and

$$d > a + ab.$$

This proves

$$\sum_{y \in \mathfrak{a}'}^{(B)} y = a + ab.$$

Since $\mathfrak{a}'$ is a subclass of A, $a + ab$ is an element of N, as we wished to prove.

If a and b are arbitrary elements of N, then there exists a subclass $\mathfrak{C}$ of A such that

$$\sum_{x \in \mathfrak{C}}^{(B)} x = a. \tag{26}$$

From (26) follows, by Theorem 17,

$$\sum_{x \in \mathfrak{C}}^{(B)} (x + xb) = a + ab.$$

Since $x + xb$ is, according to what we have just proved, an element of N, whenever x is an element of A, and b an element of N, we find that also $a + ab$ is an element of N. Now it is clear that also $a + b = (a + ab) \vee (ab + b)$ is an element of N, and Theorem 66 is proved.

If A is an arbitrary subring of B then the subclass N of B, defined in the same way as in Theorem 66, is certainly a *sublattice* of B, but not necessarily a *subring* of B. If, in the example given at the beginning of §4, we put for instance

$$a = a_1$$

$$b = \sum_{n=2}^{\infty}{}^{(A)} a_n,$$

then $a + ab$ is not representable as a join of elements of $\mathfrak{a}$ with respect to A.

If it is known, conversely, that N is a subring of B, and B is a complete Boolean ring, one can easily conclude that A is invariant in B. But this statement does not remain true if we omit the supposition regarding the completeness of B.

Let A be the Boolean ring whose elements are the finite sets of even positive integers and the complements of these sets with respect to the set of all positive integers, and let B consist of all finite sets of positive integers and their complements with respect to the set of all positive integers; the finite operations in A and B are meant set-theoretically. Then A is a subring of B, and N has the same property because it is identical with A, but A is not invariant in B.

Another form of Theorem 66 is

THEOREM 67. *Let A be an invariant subring of a Boolean ring B, and $\mathfrak{A}$ a subclass of A. If in this case $\prod^{(B)}_{x \in \mathfrak{A}} x$ exists then there exists also a subclass $\mathfrak{B}$ of A such that $\sum^{(B)}_{y \in \mathfrak{B}} y$ exists and is equal to $\prod^{(B)}_{x \in \mathfrak{A}} x$; and if $\sum^{(\mathfrak{B})}_{x \in \mathfrak{A}} x$ exists, and moreover an element a of A such that $x < a$ for $x \in \mathfrak{A}$, then there exists a subclass $\mathfrak{C}$ of A such that $\prod^{(B)}_{z \in \mathfrak{C}} z$ exists and is equal to $\sum^{(B)}_{x \in \mathfrak{A}} x$.*

PROOF. If $\prod^{(B)}_{x \in \mathfrak{A}} x$ exists, and x_0 is an element of $\mathfrak{A}$, then $\sum^{(B)}_{x \in \mathfrak{A}} (x_0 + x_0 x)$ exists too and is equal to $x_0 + \prod^{(B)}_{x \in \mathfrak{A}} x$. (Theorem 14.) Consequently we have

$$\prod_{x \in \mathfrak{A}}^{(B)} x = x_0 + \sum_{x \in \mathfrak{A}}^{(B)} (x_0 + x_0 x).$$

Hence $\prod^{(B)}_{x \in \mathfrak{A}} x$ is, by Theorem 66, an element of N.

If there exist $\sum^{(B)}_{x \in \mathfrak{A}} x$ and an element a of A such that $x < a$ for $x \in \mathfrak{A}$, then there exists, by Theorem 66, a subclass $\mathfrak{C}^*$ of A such that

$$a + \sum_{x \in \mathfrak{A}}^{(B)} x = \sum_{z \in \mathfrak{C}^*}^{(B)} z;$$

hence we have, by Theorem 15,

$$\sum_{x \in \mathfrak{A}}^{(B)} x = \prod_{z \in \mathfrak{C}^*}^{(B)} (a + z).$$

THEOREM 68.[1] *The Boolean ring $\mathfrak{N}$ of the normal ideals of a Boolean ring A is a join-extension of the Boolean ring $\mathfrak{P}$ of the principal ideals of A. Hence $\mathfrak{P}$ is invariant in $\mathfrak{N}$.*

PROOF. If $\mathfrak{b}$ is a normal ideal in A, we have obviously

$$\mathfrak{b} = \sum_{a \in \mathfrak{b}}^{(\mathfrak{N})} \mathfrak{a}(a),$$

or

$$\mathfrak{b} = \sum_{\substack{\mathfrak{a} \subset \mathfrak{b} \\ \mathfrak{a} \in \mathfrak{P}}}^{(\mathfrak{N})} \mathfrak{a}.$$

---

[1] *Editorial comment.* Theorems 68, 69, 80, 81, and some of the other theorems of this section are either partly or wholly in H. MacNeille's "Partially ordered sets," Trans. Am. Math. Soc. 42(1937), esp. Theorem 11.13 and 11.12.

THEOREM 69. (Corollary of Theorem 68.) *Any Boolean ring can be invariantly extended to a complete Boolean ring.*

THEOREM 70. *If the signs A, B, and N have the same meaning as in Theorem 66 then N is invariant in B.*

PROOF. Let $\bar{B}$ be a complete Boolean ring which is an invariant extension of B. (That such a Boolean ring $\bar{B}$ exists follows from Theorem 69.) Then A is, by Theorem 53, an invariant subring of $\bar{B}$. Further let $\bar{N}$ be the subring of $\bar{B}$ arising from A and $\bar{B}$ in the same way as N from A and B in Theorem 66. N is a subring of $\bar{N}$. If now $\mathfrak{B}$ is a subclass of N, and b an element of N such that

$$\sum_{a \in \mathfrak{B}}^{(N)} a = b \tag{27}$$

then we put

$$\sum_{a \in \mathfrak{B}}^{(\bar{B})} a = c. \tag{28}$$

($\sum_{a \in \mathfrak{B}}^{(\bar{B})} a$ exists because $\bar{B}$ is a complete Boolean ring.) c satisfies the inequality

$$c < b$$

and is an element of $\bar{N}$. Since b is an element of $\bar{N}$, the same holds also for $b + c$. Hence we have

$$b + c = \sum_{x \in \mathfrak{C}}^{(\bar{B})} x \tag{29}$$

for an appropriate subclass $\mathfrak{C}$ of A. On the other side we get from (28)

$$(b + c)a = 0$$

for $a \in \mathfrak{B}$. Consequently we have the more

$$xa = 0$$

for $x \in \mathfrak{C}$, $a \in \mathfrak{B}$. Because $\mathfrak{C}$ is a subclass of N we get from this and from (27)

$$xb = 0$$

for $x \in \mathfrak{C}$. Now from (29) follows

$$b + c = 0$$

i.e. $c = b$. By this is proved that N is an invariant subring of $\bar{B}$. Hence N is, by Theorem 55, also invariant in B.

THEOREM 71. *If the signs A, B, and N have the same meaning as in Theorem 66, and R is a join-extension of A invariant in B* (i.e. an invariant subring of B which is a join-extension of A), *then R is contained in N.*

DEFINITION 11. *If the signs A, B, and N have the same meaning as in Theorem 66, then N is called the maximal join-extension of A invariant in B.*

Definition 11 is justified by Theorem 71.

THEOREM 72. *If A is an invariant subring of a Boolean ring B, the join-extensions of A invariant in B form a complete sublattice of the lattice of all sub-*

rings of B. The zero of this sublattice is A, and its unit is the maximal join-extension of A invariant in B.

THEOREM 73. *Let A be an invariant subring of Boolean ring B, R a join-extension of A invariant in B, N the maximal join-extension of A invariant in B, and D an invariant subring of B containing N. In this case N is also the maximal join-extension of R invariant in D.*

THEOREM 74. *If $\mathfrak{a}$ is an ideal of a Boolean ring A, the maximal join-extension of $\mathfrak{a}$ invariant in A is identical with $\mathfrak{a}''$.*

PROOF. See Theorem 18.

If $\mathfrak{a}$ is an arbitrary invariant subring of A, the maximal join-extension of $\mathfrak{a}$ invariant in A is certainly a subring of $\mathfrak{a}''$ (and, by Theorem 55, even invariant in $\mathfrak{a}''$), but not necessarily identical with $\mathfrak{a}''$. Let for instance A be a *finite* Boolean ring, and $\mathfrak{a}$ a proper subring of A which contains the unit of A. Then $\mathfrak{a}$ is certainly invariant in A, and $\mathfrak{a}'' = A$, while the maximal join-extension of $\mathfrak{a}$ invariant in A is $\mathfrak{a}$ itself.

THEOREM 75. *If A is an invariant subring of a complete Boolean ring B, the maximal join-extension N of A invariant in B is itself a complete Boolean ring, and any complete invariant subring of B containing A contains N.*

PROOF. If $\mathfrak{B}$ is a subclass of N then $\sum^{(B)}_{x \in \mathfrak{B}} x$ exists and is, according to the definition of N, contained in N; hence $\sum^{(N)}_{x \in \mathfrak{B}} x$ exists and is identical with $\sum^{(B)}_{x \in B} x$. (See Theorem 47.) If we put

$$\sum_{x \in \mathfrak{B}}^{(N)} x = a$$

then $\sum^{(N)}_{x \in \mathfrak{B}} (a + x)$ exists too. Hence, by Theorem 15, there exists also $\prod^{(N)}_{x \in \mathfrak{B}} x$ and is identical with $a + \sum^{(N)}_{x \in \mathfrak{B}} (a + x)$.

If D is a complete invariant subring of B, and D contains A, then $\sum^{(D)}_{x \in \mathfrak{A}} x$ exists whenever $\mathfrak{A}$ is a subclass of A, and we have

$$\sum_{x \in \mathfrak{A}}^{(D)} x = \sum_{x \in \mathfrak{A}}^{(B)} x.$$

Thus D contains all elements of N.

DEFINITION 12. *An extension N of a Boolean ring A is said to be a minimal complete invariant extension of A, or shorter: a completion of A, if N is a complete Boolean ring invariant over A, and there is no complete proper subring of N containing A.*

THEOREM 76. *If A is an invariant subring of a complete Boolean ring B, and N the maximal join-extension of A invariant in B, then N is a minimal complete invariant extension of A.*

PROOF. N is, by Theorem 75, a complete Boolean ring and, by Theorems 61 and 66, invariant over A. If R is a complete subring of N containing A then R is, by Theorems 53 and 64, invariant in B; hence we have, by Theorem 75,

$$R \supset N$$

and consequently

$$R = N.$$

THEOREM 77. *A Boolean ring N is a minimal complete invariant extension of a Boolean ring A if and only if N is a complete Boolean ring and a join-extension of A.*

PROOF. If N is a minimal complete invariant extension of A, from Theorem 75 and Definition 12 follows that the maximal join-extension of A invariant in N cannot be different from N; hence N is a join-extension of A.

If N is complete and a join-extension of A, we see immediately that N is the maximal join-extension of A invariant in N; hence N is, by Theorem 76, a minimal complete invariant extension of A.

THEOREM 78. *If A and N are invariant subrings of a Boolean ring B, and N is a minimal complete invariant extension of A, then N is the maximal join-extension of A invariant in B.*

The proof follows easily from Theorem 77 and Definition 11.

THEOREM 79. *Any Boolean ring has at least one minimal complete invariant extension.*

PROOF. See Theorems 69 and 76.

THEOREM 80. *The Boolean ring $\mathfrak{N}$ of the normal ideals of a Boolean ring A is a minimal complete invariant extension of the Boolean ring $\mathfrak{P}$ of the principal ideals of A.*

PROOF. See Theorems 68 and 77.

THEOREM 81. *Any minimal complete invariant extension N of a Boolean ring A is isomorphic with the Boolean ring $\mathfrak{N}$ of the normal ideals of A, and there is exactly one isomorphism between N and $\mathfrak{N}$ carrying every element a of A into the principal ideal $\mathfrak{a}(a)$ of A.*

PROOF. Since N is a complete Boolean ring, every normal ideal of N is principal according to Theorem 18. If a is an element of N, $\mathfrak{b}(a)$ the principal ideal of N generated by a, and we put

$$\mathfrak{a} = A \cdot \mathfrak{b}(a) \tag{31}$$

then $\mathfrak{a}$ is a normal ideal of A, and any normal ideal of A can be obtained in this way. (Theorem 57.) On the other side N is, by Theorem 77, a join-extension of A. Therefore we get from (31)

$$a = \sum_{x \in \mathfrak{a}}^{(N)} x. \tag{32}$$

(See Theorem 58.) We see that to different elements a of N there correspond different normal ideals $\mathfrak{a}$ of $\mathfrak{N}$. Finally it is obvious that the biunivocal correspondence between N and $\mathfrak{N}$ defined by (31) or (32) preserves inclusion. Hence it is an isomorphism between N and $\mathfrak{N}$. If $a \in A$, then $A \cdot \mathfrak{b}(a) = \mathfrak{a}(a)$. That there is but one such isomorphism follows from Theorems 59 and 68.

THEOREM 82. *If A and A^* are Boolean rings between which there exists an isomorphism Λ, N a minimal complete invariant extension of A, and N^* a minimal complete invariant extension of A^*, then there exists exactly one extension of the isomorphism Λ to an isomorphism Λ_N between N and N^*.*

In particular we may say that any two minimal complete invariant extensions of the same Boolean ring are abstractly identical.

The proof of Theorem 82 follows easily from Theorems 59 and 81.

Now we can state the following generalization of Theorem 59.

THEOREM 83. *If A and A^* are Boolean rings between which there exists an isomorphism Λ, R a join-extension of A, and $\hat{R}$ a join-extension of A^*, then there exists at the utmost one subring R^* of $\hat{R}$ containing A^* such that the isomorphism Λ can be extended to an isomorphism Λ_R between R and R^*. R^* exists if and only if $\sum_{a^* \in \mathfrak{a}^*}^{(\hat{R})} a^*$ exists whenever b is an element of R, and $\mathfrak{a}^*$ is the subclass of A^* assigned to the element b of R in the same way as in Theorem 59. In this case R^* is the class of all such joins $\sum_{a^* \in \mathfrak{a}^*}^{(\hat{R})} a^*$, the isomorphism Λ_R is unique, and the element b^* of R^*, assigned to the element b of R by virtue of Λ_R, is given by the equation*

$$b^* = \sum_{a^* \in \mathfrak{a}^*}^{(\hat{R})} a^*.$$

Hence if $\hat{R}$ is especially a completion of A^*, then R^* certainly exists.

PROOF. In face of Theorems 59, 61, and 64 we may restrict ourselves to prove that the existence of all joins $\sum_{a^* \in \mathfrak{a}^*}^{(\hat{R})} a^*$ implies that R^* exists. Let N be a completion of R, and N^* a completion of $\hat{R}$. (See Theorem 79.) It is obvious that N and N^* are also completions of A and A^* respectively. Hence the isomorphism Λ can be extended to an isomorphism Λ_N between N and N^* in exactly one way. (Theorem 82.) And in face of Theorem 59 it is obvious that the existence of all our joins $\sum_{a^* \in \mathfrak{a}^*}^{(\hat{R})} a^*$ implies that the subring R^* of N^* into which R is carried by Λ_N is contained in $\hat{R}$. Thus we have proved what we had wished to.

In the rest of this paragraph we shall state some corollaries of Theorems 46, 49, 50, 51, 52, and 68. Here and in the next paragraph we shall consider a fixed Boolean ring A; the Boolean ring of the principal ideals of A will be denoted by $\mathfrak{P}$, and the Boolean ring of the normal ideals of A by $\mathfrak{N}$. Notice Theorem 38.

THEOREM 84. *An element a is an interelement of a sequence a_n if and only if*

$$(33) \qquad {}^{(\mathfrak{N})}\underline{\lim}_{n\to\infty}\, \mathfrak{a}(a_n) \subset \mathfrak{a}(a) \subset {}^{(\mathfrak{N})}\overline{\lim}_{n\to\infty}\, \mathfrak{a}(a_n).$$

PROOF. See Theorems 11, 46, 49, and 68.

THEOREM 85. *$\underline{\lim}_{n\to\infty} a_n$ exists if and only if ${}^{(\mathfrak{N})}\underline{\lim}_{n\to\infty} \mathfrak{a}(a_n)$ is a principal ideal $\mathfrak{a}(a)$; in this case we have*

$$\underline{\lim}_{n\to\infty} a_n = a.$$

PROOF. See Theorems 51 and 68.

THEOREM 86. *$\overline{\lim}_{n\to\infty} a_n$ exists if and only if ${}^{(\mathfrak{N})}\overline{\lim}_{n\to\infty} \mathfrak{a}(a_n)$ is a principal ideal $\mathfrak{a}(a)$; in this case we have*

$$\overline{\lim}_{n\to\infty} a_n = a.$$

PROOF. See Theorems 52 and 68.

THEOREM 87. *A sequence of elements a_n of a Boolean ring A is convergent with*

respect to A if and only if the sequence of the principal ideals $\mathfrak{a}(a_n)$ *is convergent with respect to* $\mathfrak{N}$, *and its limit is a principal ideal* $\mathfrak{a}(a)$; *in this case we have*

$$^{(A)}\lim_{n\to\infty} a_n = a.$$

PROOF. See Theorems 50 and 68.

At this place we mention a theorem which may be considered as a generalization of Theorem 20:

THEOREM 88. *If* $\mathfrak{U}$ *is the class of the subelements of a sequence* a_n, *we have*

$$^{(\mathfrak{N})}\lim_{n\to\infty} \mathfrak{a}(a_n) = \mathfrak{U}'' . \tag{34}$$

PROOF. It is evident that

$$\sum_{n=1}^{\infty}{}^{(\mathfrak{Z})} \prod_{k=n}^{\infty}{}^{(\mathfrak{N})} \mathfrak{a}(a_k) = \mathfrak{U}$$

if $\mathfrak{Z}$ is the distributive lattice of the ideals of the given Boolean ring, and from this and from Theorem 36 there ensues immediately the asserted Relation (34).

6. Relations between Interelements, Lower and Upper Limits, and Limits

THEOREM 89. *If* $a_n < b_n$ *for* $n = 1, 2, 3, \cdots$, a *is an interelement of the sequence* a_n, *and* $\overline{\lim}_{n\to\infty} b_n$ *exists, then*

$$a < \overline{\lim_{n\to\infty}} b_n . \tag{35}$$

THEOREM 90. *If* $a_n < b_n$ *for* $n = 1, 2, 3, \cdots$, $\underline{\lim}_{n\to\infty} a_n$ *exists, and* b *is an interelement of the sequence* b_n, *then*

$$\underline{\lim_{n\to\infty}} a_n < b. \tag{36}$$

PROOF OF THEOREMS 89 AND 90. It is obvious that $a_n < b_n$ for $n = 1, 2, 3, \cdots$ implies

$$^{(\mathfrak{N})}\underline{\lim_{n\to\infty}} \mathfrak{a}(a_n) \subset {}^{(\mathfrak{N})}\underline{\lim_{n\to\infty}} \mathfrak{a}(b_n)$$

and

$$^{(\mathfrak{N})}\overline{\lim_{n\to\infty}} \mathfrak{a}(a_n) \subset {}^{(\mathfrak{N})}\overline{\lim_{n\to\infty}} \mathfrak{a}(b_n).$$

After this is settled Theorems 89 and 90 can be easily concluded from Theorem 84.

THEOREM 91. *If* $a_n < b_n$ *for* $n = 1, 2, 3, \cdots$ *then those of the relations*

$$\underline{\lim_{n\to\infty}} a_n < \underline{\lim_{n\to\infty}} b_n$$

$$\underline{\lim_{n\to\infty}} a_n < \overline{\lim_{n\to\infty}} b_n$$

$$\overline{\lim_{n\to\infty}} a_n < \overline{\lim_{n\to\infty}} b_n$$

hold in which both sides exist.

PROOF. See Theorems 89 and 90.

THEOREM 92. *If* $\lim_{n\to\infty} b_n = 0$, *and* $a_n < b_n$ *for* $n = 1, 2, 3, \cdots$ *then also*

$$\lim_{n\to\infty} a_n = 0.$$

Notice that 0 is the only subelement and the only interelement of the sequence a_n.

LEMMA 8. *If* $a_k < a_l$ *and* $b_k < b_l$ *for* $k < l$, $k, l = 1, 2, 3, \cdots$ *and* $\sum_{n=1}^{\infty} a_n$ *and* $\sum_{n=1}^{\infty} b_n$ *exist, then* $\sum_{n=1}^{\infty} a_n b_n$ *exists too and is equal to* $\sum_{n=1}^{\infty} a_n \cdot \sum_{n=1}^{\infty} b_n$.

PROOF. We have obviously

$$\sum_{n=1}^{\infty} a_n \cdot \sum_{n=1}^{\infty} b_n > a_k b_k$$

for every k. If $c > a_n b_n$ for every n then

$$c > a_k b_l$$

for $k, l = 1, 2, 3, \cdots$. From this inequality we get by using twice Theorem 13 the inequality

$$c > \sum_{n=1}^{\infty} a_n \cdot \sum_{n=1}^{\infty} b_n$$

which proves the assertion.

LEMMA 9. *If* $a_k > a_l$ *and* $b_k > b_l$ *for* $k < l$, $k, l = 1, 2, 3, \cdots$ *and* $\prod_{n=1}^{\infty} a_n$ *and* $\prod_{n=1}^{\infty} b_n$ *exist, then* $\prod_{n=1}^{\infty} (a_n \vee b_n)$ *exists too and is equal to* $\prod_{n=1}^{\infty} a_n \vee \prod_{n=1}^{\infty} b_n$.

PROOF. $\prod_{n=1}^{\infty} a_n \vee \prod_{n=1}^{\infty} b_n < a_k \vee b_k$ for every k. If $c < a_n \vee b_n$ for every n then $c < a_k \vee b_l$ for $k, l = 1, 2, 3, \cdots$. Hence we get by using twice Theorem 12

$$c < \prod_{n=1}^{\infty} a_n \vee \prod_{n=1}^{\infty} b_n.$$

This proves the assertion.

THEOREM 93. *If* $\varliminf_{n\to\infty} a_n$ *and* $\varliminf_{n\to\infty} b_n$ *exist then* $\varliminf_{n\to\infty} a_n b_n$ *exists too and is equal to* $\varliminf_{n\to\infty} a_n \cdot \varliminf_{n\to\mathfrak{N}} b_n$.

PROOF.

$$\varliminf_{n\to\infty}{}^{(\mathfrak{N})}\, \mathfrak{a}(a_n b_n) = \varliminf_{n\to\infty}{}^{(\mathfrak{N})}\, \mathfrak{a}(a_n)\mathfrak{a}(b_n) = \sum_{n=1}^{\infty}{}^{(\mathfrak{N})} \prod_{k=n}^{\infty}{}^{(\mathfrak{N})}\, \mathfrak{a}(a_k)\mathfrak{a}(b_k)$$

$$= \sum_{n=1}^{\infty}{}^{(\mathfrak{N})} \left[\prod_{k=n}^{\infty}{}^{(\mathfrak{N})}\, \mathfrak{a}(a_k) \cdot \prod_{k=n}^{\infty}{}^{(\mathfrak{N})}\, \mathfrak{a}(b_k)\right];$$

hence, by Lemma 8,

$$\varliminf_{n\to\infty}{}^{(\mathfrak{N})}\, \mathfrak{a}(a_n b_n) = \sum_{n=1}^{\infty}{}^{(\mathfrak{N})} \prod_{k=n}^{\infty}{}^{(\mathfrak{N})}\, \mathfrak{a}(a_k) \cdot \sum_{n=1}^{\infty}{}^{(\mathfrak{N})} \prod_{k=n}^{\infty}{}^{(\mathfrak{N})}\, \mathfrak{a}(b_k) = \varliminf_{n\to\infty}{}^{(\mathfrak{N})}\, \mathfrak{a}(a_n) \cdot \varliminf_{n\to\infty}{}^{(\mathfrak{N})}\, \mathfrak{a}(b_n),$$

and, by Theorem 85,

$$^{(\mathfrak{R})}\varliminf_{n\to\infty} \mathfrak{a}(a_n b_n) = \mathfrak{a}(\varliminf_{n\to\infty} a_n)\cdot\mathfrak{a}(\varliminf_{n\to\infty} b_n) = \mathfrak{a}(\varliminf_{n\to\infty} a_n \cdot \varliminf_{n\to\infty} b_n);$$

from this follows, again by Theorem 85, the assertion.

THEOREM 94. *If* $\varlimsup_{n\to\infty} a_n$ *and* $\varlimsup_{n\to\infty} b_n$ *exist then* $\varlimsup_{n\to\infty} (a_n \vee b_n)$ *exists too and is equal to* $\varlimsup_{n\to\infty} a_n \vee \varlimsup_{n\to\infty} b_n$.

PROOF. Compare the proof of Theorem 93 and see Lemma 9 and Theorem 86.

THEOREM 95. *If* $\varliminf_{n\to\infty} a_n$ *and* $\varlimsup_{n\to\infty} b_n$ *exist then* $\varliminf_{n\to\infty} (a_n + a_n b_n)$ *exists too and is equal to* $\varliminf_{n\to\infty} a_n + \varliminf_{n\to\infty} a_n \cdot \varlimsup_{n\to\infty} b_n$.

PROOF. We have

$$\mathfrak{a}(a_n + a_n b_n) = \mathfrak{a}(a_n)\cdot\mathfrak{a}'(b_n)$$

and, by Lemma 6,

$$^{(\mathfrak{R})}\varliminf_{n\to\infty} \mathfrak{a}'(b_n) = [^{(\mathfrak{R})}\varlimsup_{n\to\infty} \mathfrak{a}(b_n)]';$$

hence we have, by Theorem 93,

$$^{(\mathfrak{R})}\varliminf_{n\to\infty} \mathfrak{a}(a_n + a_n b_n) = \mathfrak{a}(\varliminf_{n\to\infty} a_n)\cdot\mathfrak{a}'(\varlimsup_{n\to\infty} b_n) = \mathfrak{a}(\varliminf_{n\to\infty} a_n + \varliminf_{n\to\infty} a_n \cdot \varlimsup_{n\to\infty} b_n).$$

THEOREM 96. *If a is an interelement of the sequence a_n, and b an interelement of the sequence b_n, then those of the inequalities*

$$ab > \varliminf_{n\to\infty} a_n b_n,$$

$$a \vee b < \varlimsup_{n\to\infty} (a_n \vee b_n),$$

and

$$a + ab > \varliminf_{n\to\infty} (a_n + a_n b_n)$$

hold the right-hand sides of which exist.

PROOF. See Theorems 84, 93, 94, and 95.

THEOREM 97. *If a is an interelement of the sequence a_n, and* $\varliminf_{n\to\infty} b_n$ *exists, then* $a \varliminf_{n\to\infty} b_n$ *is an interelement of the sequence $a_n b_n$.*

PROOF. We have for $l = 1, 2, 3, \cdots$

$$^{(\mathfrak{R})}\varlimsup_{n\to\infty} \mathfrak{a}(a_n b_n) = {}^{(\mathfrak{R})}\varlimsup_{n\to\infty} [\mathfrak{a}(a_n)\mathfrak{a}(b_n)] = \prod_{n=1}^{\infty}{}^{(\mathfrak{R})} \sum_{k=n}^{\infty}{}^{(\mathfrak{R})} \mathfrak{a}(a_k)\mathfrak{a}(b_k)$$

$$= \prod_{n=l}^{\infty}{}^{(\mathfrak{R})} \sum_{k=n}^{\infty}{}^{(\mathfrak{R})} \mathfrak{a}(a_k)\mathfrak{a}(b_k) \supset \left[\prod_{n=l}^{\infty}{}^{(\mathfrak{R})} \sum_{k=n}^{\infty}{}^{(\mathfrak{R})} \mathfrak{a}(a_k)\right]\cdot\prod_{m=l}^{\infty}{}^{(\mathfrak{R})} \mathfrak{a}(b_m)$$

$$= {}^{(\mathfrak{R})}\varlimsup_{n\to\infty} \mathfrak{a}(a_n)\cdot\prod_{m=l}^{\infty}{}^{(\mathfrak{R})} \mathfrak{a}(b_m),$$

hence

$$^{(\mathfrak{N})}\overline{\lim_{n\to\infty}}\,\mathfrak{a}(a_nb_n) \supset {}^{(\mathfrak{N})}\overline{\lim_{n\to\infty}}\,\mathfrak{a}(a_n)\cdot{}^{(\mathfrak{N})}\underline{\lim_{n\to\infty}}\,\mathfrak{a}(b_n) = {}^{(\mathfrak{N})}\overline{\lim_{n\to\infty}}\,\mathfrak{a}(a_n)\cdot\mathfrak{a}(\underline{\lim_{n\to\infty}}\,b_n)$$

$$\supset \mathfrak{a}(a)\cdot(\underline{\lim_{n\to\infty}}\,b_n) = \mathfrak{a}(a\,\underline{\lim_{n\to\infty}}\,b_n).$$

On the other side it follows from the first relation of Theorem 96 that

$$\mathfrak{a}(a)\cdot\mathfrak{a}(\underline{\lim_{n\to\infty}}\,b_n) \supset {}^{(\mathfrak{N})}\underline{\lim_{n\to\infty}}\,[\mathfrak{a}(a_n)\mathfrak{a}(b_n)],$$

or

$$\mathfrak{a}(a\,\underline{\lim_{n\to\infty}}\,b_n) \supset {}^{(\mathfrak{N})}\underline{\lim_{n\to\infty}}\,\mathfrak{a}(a_nb_n).$$

This proves the assertion, according to Theorem 84.

THEOREM 98. (Corollary of Theorem 97.) *If* $\overline{\lim_{n\to\infty}}\,a_n$ *and* $\underline{\lim_{n\to\infty}}\,b_n$ *exist then* $\overline{\lim_{n\to\infty}}\,a_n\cdot\underline{\lim_{n\to\infty}}\,b_n$ *is an interelement of the sequence* a_nb_n.

THEOREM 99. *If* a *is an interelement of the sequence* a_n, *and* $\overline{\lim_{n\to\infty}}\,b_n$ *exists, then* $a \vee \overline{\lim_{n\to\infty}}\,b_n$ *is an interelement of the sequence* $a_n \vee b_n$.

Theorem 99 can be proved in a way dually corresponding to the proof of Theorem 97.

THEOREM 100. (Corollary of Theorem 99.) *If* $\underline{\lim_{n\to\infty}}\,a_n$ *and* $\overline{\lim_{n\to\infty}}\,b_n$ *exist then* $\underline{\lim_{n\to\infty}}\,a_n \vee \overline{\lim_{n\to\infty}}\,b_n$ *is an interelement of the sequence* $a_n \vee b_n$.

THEOREM 101. *If* a *is an interelement of the sequence* a_n, *and* $\overline{\lim_{n\to\infty}}\,b_n$ *exists, then* $a + a\,\overline{\lim_{n\to\infty}}\,b_n$ *is an interelement of the sequence* $a_n + a_nb_n$.

PROOF. We have

$$\mathfrak{a}(a + a\,\overline{\lim_{n\to\infty}}\,b_n) = \mathfrak{a}(a)\cdot\overline{\lim_{n\to\infty}}\,\mathfrak{a}'(b_n)$$

(see Lemma 6); hence $\mathfrak{a}(a + a\,\overline{\lim_{n\to\infty}}\,b_n)$ is, by Theorem 97, an interelement of the sequence $\mathfrak{a}(a_n)\mathfrak{a}'(b_n) = \mathfrak{a}(a_n + a_nb_n)$ with respect to $\mathfrak{N}$.

THEOREM 102. (Corollary of Theorem 101.) *If* $\overline{\lim_{n\to\infty}}\,a_n$ *and* $\overline{\lim_{n\to\infty}}\,b_n$ *exist then* $\overline{\lim_{n\to\infty}}\,a_n + \overline{\lim_{n\to\infty}}\,a_n\cdot\overline{\lim_{n\to\infty}}\,b_n$ *is an interelement of the sequence* $a_n + a_nb_n$.

THEOREM 103. *If* $\underline{\lim_{n\to\infty}}\,a_n$ *exists, and* b *is an interelement of the sequence* b_n, *then* $\underline{\lim_{n\to\infty}}\,a_n + b\,\underline{\lim_{n\to\infty}}\,a_n$ *is an interelement of the sequence* $a_n + a_nb_n$.

The proof follows from the equation

$$\mathfrak{a}(\underline{\lim_{n\to\infty}}\,a_n + b\,\underline{\lim_{n\to\infty}}\,a_n) = {}^{(\mathfrak{N})}\underline{\lim_{n\to\infty}}\,\mathfrak{a}(a_n)\cdot\mathfrak{a}'(b)$$

and from Theorem 97; for $\mathfrak{a}'(b)$ is, by Lemma 7, an interelement of the sequence $\mathfrak{a}'(b_n)$ with respect to $\mathfrak{N}$.

THEOREM 104. (Corollary of Theorem 103.) *If* $\underline{\lim}_{n\to\infty} a_n$ *and* $\underline{\lim}_{n\to\infty} b_n$ *exist then* $\underline{\lim}_{n\to\infty} a_n + \underline{\lim}_{n\to\infty} a_n \cdot \underline{\lim}_{n\to\infty} b_n$ *is an interelement of the sequence* $a_n + a_n b_n$.

THEOREM 105. *If* $\overline{\lim}_{n\to\infty} a_n$ *and* $\overline{\lim}_{n\to\infty} b_n$ *exist, and* x *is an interelement of the sequence* $a_n b_n$, *then*

$$x < \overline{\lim}_{n\to\infty} a_n \cdot \overline{\lim}_{n\to\infty} b_n.$$

PROOF. We have, by Theorem 89,

$$x < \overline{\lim}_{n\to\infty} a_n$$

and

$$x < \overline{\lim}_{n\to\infty} b_n;$$

hence

$$x < \overline{\lim}_{n\to\infty} a_n \cdot \overline{\lim}_{n\to\infty} b_n.$$

THEOREM 106. *If* $\underline{\lim}_{n\to\infty} a_n$ *and* $\underline{\lim}_{n\to\infty} b_n$ *exist, and* x *is an interelement of the sequence* $a_n \vee b_n$, *then*

$$x > \underline{\lim}_{n\to\infty} a_n \vee \underline{\lim}_{n\to\infty} b_n.$$

PROOF. We have, by Theorem 90,

$$x > \underline{\lim}_{n\to\infty} a_n$$

and

$$x > \underline{\lim}_{n\to\infty} b_n;$$

hence

$$x > \underline{\lim}_{n\to\infty} a_n \vee \underline{\lim}_{n\to\infty} b_n.$$

THEOREM 107. *If* $\overline{\lim}_{n\to\infty} a_n$ *and* $\underline{\lim}_{n\to\infty} b_n$ *exist, and* x *is an interelement of the sequence* $a_n + a_n b_n$, *then*

$$x < \overline{\lim}_{n\to\infty} a_n + \overline{\lim}_{n\to\infty} a_n \cdot \underline{\lim}_{n\to\infty} b_n.$$

PROOF. Since $\mathfrak{a}(x)$ is an interelement of the sequence $\mathfrak{a}(a_n + a_n b_n) = \mathfrak{a}(a_n)\mathfrak{a}'(b_n)$ with respect to $\mathfrak{R}$, we have, by Theorem 105 and Lemma 6,

$$\mathfrak{a}(x) \subset {}^{(\mathfrak{R})}\overline{\lim}_{n\to\infty} \mathfrak{a}(a_n) \cdot {}^{(\mathfrak{R})}\overline{\lim}_{n\to\infty} \mathfrak{a}'(b_n) = {}^{(\mathfrak{R})}\overline{\lim}_{n\to\infty} \mathfrak{a}(a_n) \cdot ({}^{(\mathfrak{R})}\underline{\lim}_{n\to\infty} \mathfrak{a}(b_n))'$$

$$= \mathfrak{a}(\overline{\lim}_{n\to\infty} a_n) \cdot \mathfrak{a}'(\underline{\lim}_{n\to\infty} b_n) = \mathfrak{a}(\overline{\lim}_{n\to\infty} a_n + \overline{\lim}_{n\to\infty} a_n \cdot \underline{\lim}_{n\to\infty} b_n).$$

THEOREM 108. *Those of the following inclusion-relations* (37) *till* (53) *which have a sense are valid:*

$$\underline{\lim}_{n\to\infty} a_n b_n < \overline{\lim}_{n\to\infty} a_n \cdot \underline{\lim}_{n\to\infty} b_n \tag{37}$$

$$\underline{\lim}_{n\to\infty} a_n b_n < \overline{\lim}_{n\to\infty} a_n \cdot \overline{\lim}_{n\to\infty} b_n \tag{38}$$

$$\overline{\lim}_{n\to\infty} a_n b_n > \overline{\lim}_{n\to\infty} a_n \cdot \underline{\lim}_{n\to\infty} b_n \tag{39}$$

$$\overline{\lim}_{n\to\infty} a_n b_n < \overline{\lim}_{n\to\infty} a_n \cdot \overline{\lim}_{n\to\infty} b_n \tag{40}$$

$$\overline{\lim}_{n\to\infty} (a_n \vee b_n) > \underline{\lim}_{n\to\infty} a_n \vee \overline{\lim}_{n\to\infty} b_n \tag{41}$$

$$\overline{\lim}_{n\to\infty} (a_n \vee b_n) > \underline{\lim}_{n\to\infty} a_n \vee \underline{\lim}_{n\to\infty} b_n \tag{42}$$

$$\underline{\lim}_{n\to\infty} (a_n \vee b_n) < \underline{\lim}_{n\to\infty} a_n \vee \overline{\lim}_{n\to\infty} b_n \tag{43}$$

$$\underline{\lim}_{n\to\infty} (a_n \vee b_n) > \underline{\lim}_{n\to\infty} a_n \vee \underline{\lim}_{n\to\infty} b_n \tag{44}$$

$$\underline{\lim}_{n\to\infty} (a_n + a_n b_n) < \underline{\lim}_{n\to\infty} a_n + \underline{\lim}_{n\to\infty} a_n \cdot \underline{\lim}_{n\to\infty} b_n \tag{45}$$

$$\underline{\lim}_{n\to\infty} (a_n + a_n b_n) < \overline{\lim}_{n\to\infty} a_n + \overline{\lim}_{n\to\infty} a_n \cdot \overline{\lim}_{n\to\infty} b_n \tag{46}$$

$$\underline{\lim}_{n\to\infty} (a_n + a_n b_n) < \overline{\lim}_{n\to\infty} a_n + \overline{\lim}_{n\to\infty} a_n \cdot \underline{\lim}_{n\to\infty} b_n \tag{47}$$

$$\overline{\lim}_{n\to\infty} (a_n + a_n b_n) > \underline{\lim}_{n\to\infty} a_n + \underline{\lim}_{n\to\infty} a_n \cdot \underline{\lim}_{n\to\infty} b_n \tag{48}$$

$$\overline{\lim}_{n\to\infty} (a_n + a_n b_n) > \overline{\lim}_{n\to\infty} a_n + \overline{\lim}_{n\to\infty} a_n \cdot \overline{\lim}_{n\to\infty} b_n \tag{49}$$

$$\overline{\lim}_{n\to\infty} (a_n + a_n b_n) < \overline{\lim}_{n\to\infty} a_n + \overline{\lim}_{n\to\infty} a_n \cdot \underline{\lim}_{n\to\infty} b_n \tag{50}$$

$$\underline{\lim}_{n\to\infty} (a_n + b_n) < \underline{\lim}_{n\to\infty} a_n + \overline{\lim}_{n\to\infty} b_n \tag{51}$$

$$\overline{\lim}_{n\to\infty} (a_n + b_n) > \underline{\lim}_{n\to\infty} a_n + \underline{\lim}_{n\to\infty} b_n \tag{52}$$

$$\overline{\lim}_{n\to\infty} (a_n + b_n) > \overline{\lim}_{n\to\infty} a_n + \overline{\lim}_{n\to\infty} b_n \tag{53}$$

PROOF.

Relation	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)	(46)	(47)	(48)	(49)	(50)
Follows from Theorem	98	105	98	105	100	106	100	106	104	102	107	104	102	107

If $\underset{n\to\infty}{\lim} a_n$, $\overline{\underset{n\to\infty}{\lim}} b_n$, and $\underset{n\to\infty}{\lim} (a_n + b_n)$ exist then $\underset{n\to\infty}{\lim} (a_n + a_n b_n)$ exists too and is equal to $\underset{n\to\infty}{\lim} a_n + \underset{n\to\infty}{\lim} a_n \cdot \overline{\underset{n\to\infty}{\lim}} b_n$. (Theorem 95.) On the other side we have $a_n + a_n b_n = (a_n + b_n)a_n$ whence

$$\underset{n\to\infty}{\lim} (a_n + a_n b_n) = \underset{n\to\infty}{\lim} (a_n + b_n) \cdot \underset{n\to\infty}{\lim} a_n .$$

(Theorem 93.) Thus we have

$$\underset{n\to\infty}{\lim} (a_n + b_n) \underset{n\to\infty}{\lim} a_n = \underset{n\to\infty}{\lim} a_n + \underset{n\to\infty}{\lim} a_n \cdot \overline{\underset{n\to\infty}{\lim}} b_n . \tag{54}$$

Further $\underset{n\to\infty}{\lim} (a_n + b_n) \overline{\underset{n\to\infty}{\lim}} b_n$ is, by Theorem 98, an interelement of the sequence $(a_n + b_n)b_n = a_n b_n + b_n$. Hence we have, by Theorem 107,

$$\underset{n\to\infty}{\lim} (a_n + b_n) \overline{\underset{n\to\infty}{\lim}} b_n < \underset{n\to\infty}{\lim} a_n \cdot \overline{\underset{n\to\infty}{\lim}} b_n + \overline{\underset{n\to\infty}{\lim}} b_n . \tag{55}$$

From (54) and (55) we get

$$\underset{n\to\infty}{\lim} (a_n + b_n) \cdot \left(\underset{n\to\infty}{\lim} a_n \vee \overline{\underset{n\to\infty}{\lim}} b_n\right) < \underset{n\to\infty}{\lim} a_n + \overline{\underset{n\to\infty}{\lim}} b_n . \tag{56}$$

Since $\underset{n\to\infty}{\lim} a_n \vee \overline{\underset{n\to\infty}{\lim}} b_n$ is, by Theorem 100, an interelement of the sequence $a_n \vee b_n$, and $a_n + b_n < a_n \vee b_n$, we have, by Theorem 90,

$$\underset{n\to\infty}{\lim} (a_n + b_n) < \underset{n\to\infty}{\lim} a_n \vee \overline{\underset{n\to\infty}{\lim}} b_n .$$

Thus (56) involves (51).

If $\underset{n\to\infty}{\lim} a_n$, $\underset{n\to\infty}{\lim} b_n$, and $\overline{\underset{n\to\infty}{\lim}} (a_n + b_n)$ exist then $\underset{n\to\infty}{\lim} a_n + \underset{n\to\infty}{\lim} a_n \cdot \underset{n\to\infty}{\lim} b_n$ and $\underset{n\to\infty}{\lim} a_n \cdot \underset{n\to\infty}{\lim} b_n + \underset{n\to\infty}{\lim} b_n$ are, by Theorem 104, interelements of the sequences $a_n + a_n b_n$ and $a_n b_n + b_n$ respectively. Hence we have, according to the second relation of Theorem 96,

$$\left(\underset{n\to\infty}{\lim} a_n + \underset{n\to\infty}{\lim} a_n \cdot \underset{n\to\infty}{\lim} b_n\right) \vee \left(\underset{n\to\infty}{\lim} a_n \cdot \underset{n\to\infty}{\lim} b_n + \underset{n\to\infty}{\lim} b_n\right) < \overline{\underset{n\to\infty}{\lim}} \left[(a_n + a_n b_n) \vee (a_n b_n + b_n)\right],$$

i.e. Relation (52). Relation (53) we can prove in the same way, using Theorem 102 instead of Theorem 104.

THEOREM 109. *If* $\overline{\underset{n\to\infty}{\lim}} a_n$ *and* $\underset{n\to\infty}{\lim} b_n$ *exist then* $\overline{\underset{n\to\infty}{\lim}} a_n b_n$ *exists too and is equal to* $\overline{\underset{n\to\infty}{\lim}} a_n \cdot \underset{n\to\infty}{\lim} b_n$.

PROOF. See Theorems 98 and 105.

THEOREM 110. *If* $\underset{n\to\infty}{\lim} a_n$ *and* $\underset{n\to\infty}{\lim} b_n$ *exist then* $\underset{n\to\infty}{\lim} (a_n \vee b_n)$ *exists too and is equal to* $\underset{n\to\infty}{\lim} a_n \vee \underset{n\to\infty}{\lim} b_n$.

PROOF. See Theorems 100 and 106.

THEOREM 111. *If* $\lim_{n\to\infty} a_n$ *and* $\underline{\lim}_{n\to\infty} b_n$ *exist then* $\overline{\lim}_{n\to\infty} (a_n + a_n b_n)$ *exists too and* is equal to $\lim_{n\to\infty} a_n + \lim_{n\to\infty} a_n \cdot \underline{\lim}_{n\to\infty} b_n$.

PROOF. See Theorems 104 and 107.

THEOREM 112. *If* $\overline{\lim}_{n\to\infty} a_n$ *and* $\lim_{n\to\infty} b_n$ *exist then* $\overline{\lim}_{n\to\infty} (a_n + a_n b_n)$ *exists too and is equal to* $\overline{\lim}_{n\to\infty} a_n + \overline{\lim}_{n\to\infty} a_n \cdot \lim_{n\to\infty} b_n$.

PROOF. See Theorems 102 and 107.

THEOREM 113. *If* $\underline{\lim}_{n\to\infty} a_n$ *and* $\lim_{n\to\infty} b_n$ *exists then* $\underline{\lim}_{n\to\infty} a_n + \lim_{n\to\infty} b_n$ *is an inter-element of the sequence* $a_n + b_n$.

PROOF. We have, by Theorem 95,

$$\underline{\lim}_{n\to\infty} (a_n + a_n b_n) = \underline{\lim}_{n\to\infty} a_n + \underline{\lim}_{n\to\infty} a_n \cdot \lim_{n\to\infty} b_n \tag{57}$$

and, by Theorem 111,

$$\overline{\lim}_{n\to\infty} (a_n b_n + b_n) = \underline{\lim}_{n\to\infty} a_n \cdot \lim_{n\to\infty} b_n + \lim_{n\to\infty} b_n . \tag{58}$$

From (57) and (58) and from Theorem 100 follows Theorem 113.

THEOREM 114. *If* $\lim_{n\to\infty} a_n$ *and* $\overline{\lim}_{n\to\infty} b_n$ *exist then* $\lim_{n\to\infty} a_n + \overline{\lim}_{n\to\infty} b_n$ *is an inter-element of the sequence* $a_n + b_n$.

PROOF. Compare the proof of Theorem 113 and see Theorems 95, 112, and 100.

THEOREM 115. *If the sequences* a_n *and* b_n *converge, the sequences* $a_n b_n$, $a_n \vee b_n$, *and* $a_n + b_n$ *converge too, and we have*

$$\lim_{n\to\infty} a_n b_n = \lim_{n\to\infty} a_n \cdot \lim_{n\to\infty} b_n , \tag{59}$$

$$\lim_{n\to\infty} (a_n \vee b_n) = \lim_{n\to\infty} a_n \vee \lim_{n\to\infty} b_n , \text{ and} \tag{60}$$

$$\lim_{n\to\infty} (a_n + b_n) = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n . \tag{61}$$

PROOF. (59) follows from Theorems 93 and 109, and (60) from Theorems 94 and 110. Further Theorems 95 and 111 (or 112) involve

$$\lim_{n\to\infty} (a_n + a_n b_n) = \lim_{n\to\infty} a_n + \lim_{n\to\infty} a_n \cdot \lim_{n\to\infty} b_n \tag{62}$$

and

$$\lim_{n\to\infty} (a_n b_n + b_n) = \lim_{n\to\infty} a_n \cdot \lim_{n\to\infty} b_n + \lim_{n\to\infty} b_n . \tag{63}$$

From (62) and (63) follows (61), by the aid of (60).

THEOREM 116. *If* $a_k a_l = 0$ *for* $k, l = 1, 2, 3, \cdots, k \neq l$ *then*

$$\lim_{n\to\infty} a_n = 0.$$

PROOF. 0 is obviously the only subelement of the sequence a_n ; hence we have, by Theorem 21,

$$\lim_{n\to\infty} a_n = 0.$$

If a is an arbitrary interelement of the sequence a_n , then aa_k is an interelement of the sequence $a_n a_k$ for every k. Hence $aa_k = 0$ for $k = 1, 2, 3, \cdots$. (See Theorem 25.) Since a is also an interelement of the sequence aa_n , we obtain

$$a = 0.$$

This proves the theorem.

7. The Closure-Topology $\Gamma(\tau(A))$

The sequential topology $\tau(A)$ in a Boolean ring A, defined by Definition 8, determines a closure-topology $\Gamma(\tau(A))$. (Compare (III), Theorem 15.) Our Theorem 29 involves that a subclass of A is closed if and only if it contains with every convergent sequence also its limit.

THEOREM 117. *Any closed subring* $\mathfrak{a}$ *of a Boolean* σ*-ring* A *is itself a Boolean* σ*-ring.*

PROOF. If a_n $(n = 1, 2, 3, \cdots)$ are elements of $\mathfrak{a}$ then $\prod_{n=1}^{\infty}{}^{(A)} a_n$ and $\sum_{n=1}^{\infty}{}^{(A)} a_n$ exist and are contained in $\mathfrak{a}$. Hence, by Theorem 47, $\prod_{n=1}^{\infty}{}^{(\mathfrak{a})} a_n$ and $\sum_{n=1}^{\infty}{}^{(\mathfrak{a})} a_n$ exist too. (They are even identical with $\prod_{n=1}^{\infty}{}^{(A)} a_n$ and $\sum_{n=1}^{\infty}{}^{(A)} a_n$ respectively.)

THEOREM 118. *If* $\mathfrak{a}$ *is an invariant* σ*-subring of a Boolean ring* A*, and* a_n $(n = 1, 2, 3, \cdots)$ *are elements of* $\mathfrak{a}$*, then* ${}^{(A)}\lim_{n\to\infty} a_n$ *and* ${}^{(A)}\overline{\lim}_{n\to\infty} a_n$ *exist and are contained in* $\mathfrak{a}$*. Hence* $\mathfrak{a}$ *is especially a closed subring of* A*.*

PROOF. See Theorems 38, 51, and 52.

THEOREM 119. *If* a_n $(n = 1, 2, 3, \cdots)$ *are elements of a closed ideal* $\mathfrak{a}$ *of a Boolean ring* A*, and* a *is an interelement of the sequence* a_n *with respect to* A*, then* a *is contained in* $\mathfrak{a}$*.*

PROOF. We have, by Theorem 42,

$${}^{(A)}\lim_{n\to\infty} \sum_{k=1}^{n}{}^{(A)} a_k a = \sum_{k=1}^{\infty}{}^{(A)} a_k a = a.$$

$a_k a$ $(k = 1, 2, 3, \cdots)$ are elements of $\mathfrak{a}$ because $\mathfrak{a}$ is an ideal. The same holds also for all elements of the sequence $\sum_{k=1}^{n}{}^{(A)} a_k a$ $(n = 1, 2, 3, \cdots)$; thus also for its limit a because $\mathfrak{a}$ is closed.

THEOREM 120. *Every normal ideal of a Boolean ring is closed.*

Theorem 120 can be easily deduced from the definition of a normal ideal and from Lemma 5.

THEOREM 121. *If A is an invariant subring of a Boolean ring B, N is the maximal join-extension of A invariant in B, a_n $(n = 1, 2, 3, \cdots)$ are elements of N, a is an element of B, and either of the relations* $^{(B)}\varliminf_{n\to\infty} a_n = a$ *and* $^{(B)}\varlimsup_{n\to\infty} a_n = a$ *is valid, then a is contained in N. N is especially a closed subring of B.*

PROOF. Let $\bar{B}$ be a complete invariant extension of B (see Theorem 69) so that A is, by Theorem 53, invariant in $\bar{B}$, and $\bar{N}$ the maximal join-extension of A invariant in $\bar{B}$. (Compare the proof of Theorem 70.) If now a_n $(n = 1, 2, 3, \cdots)$ are elements of N, and $^{(B)}\varliminf_{n\to\infty} a_n$ exists, we have, by Theorem 51,

$$^{(\bar{B})}\varliminf_{n\to\infty} a_n = {}^{(B)}\varliminf_{n\to\infty} a_n ;$$

on the other side the elements a_n are also elements of $\bar{N}$. Hence $^{(B)}\varliminf_{n\to\infty} a_n$ is, by Theorems 70, 75, and 118, an element of $\bar{N}$. From this and from the definitions of N and $\bar{N}$ it can be easily concluded that $^{(B)}\varliminf_{n\to\infty} a_n$ is an element of N. If a_n $(n = 1, 2, 3, \cdots)$ are elements of N, and $^{(B)}\varlimsup_{n\to\infty} a_n$ exists, we can conclude in a similar way.

THEOREM 122. *If A is an invariant subring of a Boolean σ-ring B, the maximal join-extension of A invariant in B is itself a Boolean σ-ring.*

PROOF. See Theorems 117 and 121.

8. Double Sequences

DEFINITION 13. *Suppose that to every element of some class Θ of ordered couples (m, n) of natural numbers there is assigned an element a_{mn} of a given Boolean ring A. We say that such a correspondence defines a double sequence in A if to every natural number k there exist two natural numbers m and n such that $m \geqq k$, $n \geqq k$, and $(m, n) \in \Theta$.*

DEFINITION 14. *Let $a^{(1)}_{mn}$ and $a^{(2)}_{mn}$ be two double sequences in a Boolean ring, and $\Theta^{(1)}$ and $\Theta^{(2)}$ the classes of ordered couples of natural numbers for which they are defined. In this case we say the double sequence $a^{(1)}_{mn}$ to be a partial sequence of the the double sequence $a^{(2)}_{mn}$ if $\Theta^{(1)} \subset \Theta^{(2)}$ and $a^{(1)}_{mn} = a^{(2)}_{mn}$ for $(m, n) \in \Theta^{(1)}$.*

DEFINITION 15. *Let a_{mn} be a double sequence in a Boolean ring A, Θ the class of ordered couples of natural numbers belonging to it, and u an element of A. In this case we call u a subelement of the double sequence a_{mn} if there is a natural number k such that*

$$u < a_{mn}$$

for $m \geqq k$, $n \geqq k$, $(m, n) \in \Theta$.

DEFINITION 16. *If the signs A, a_{mn}, and Θ have the same meaning as in Definition 15, we call an element o of A a superelement of the double sequence a_{mn} if there is a natural number k such that*

$$o > a_{mn}$$

for $m \geqq k$, $n \geqq k$, $(m, n) \in \Theta$.

Now the notions of interelement, lower limit, upper limit, convergence, and limit can be defined for double sequences just so as for simple sequences. If a_{mn} is a double sequence in a Boolean ring A, the sign Θ has the same meaning as in Definition 13, and the lower limit of the double sequence a_{mn} with respect to A exists, we denote this lower limit by $\underset{\substack{m,n\to\infty\\(m,n)\,\epsilon\,\Theta}}{{}^{(A)}\underline{\lim}}\, a_{mn}$ or more simply by $\underset{m,n\to\infty}{{}^{(A)}\lim}\, a_{mn}$ in case Θ is the class of all ordered couples of natural numbers. The signs $\underset{\substack{m,n\to\infty\\(m,n)\,\epsilon\,\Theta}}{{}^{(A)}\overline{\lim}}\, a_{mn}$, $\underset{m,n\to\infty}{{}^{(A)}\overline{\lim}}\, a_{mn}$, $\underset{\substack{m,n\to\infty\\(m,n)\,\epsilon\,\Theta}}{{}^{(A)}\lim}\, a_{mn}$ and $\underset{m,n\to\infty}{{}^{(A)}\lim}\, a_{mn}$ are to be understood in similar sense. In all these signs the superscript (A) may be omitted if there is no doubt to which Boolean ring they refer.

All theorems on simple sequences obtained until now, except Theorems 31, 32, 33, 34, and 35, can be pronounced with little changes of the wording also for doubles sequences. For instance Theorems 29 and 30 assume the following forms:

THEOREM 29′: *If $a^{(1)}_{mn}$ is a partial sequence of a double sequence $a^{(2)}_{mn}$, $\Theta^{(1)}$ and $\Theta^{(2)}$ are the corresponding classes of ordered couples of natural numbers, and* $\lim\limits_{\substack{m,n\to\infty\\(m,n)\,\epsilon\,\Theta^{(2)}}} a^{(2)}_{mn}$ *exists, then* $\lim\limits_{\substack{m,n\to\infty\\(m,n)\,\epsilon\,\Theta^{(1)}}} a^{(1)}_{mn}$ *exists too and is equal to* $\lim\limits_{\substack{m,n\to\infty\\(m,n)\,\epsilon\,\Theta^{(2)}}} a^{(2)}_{mn}$.

THEOREM 30′. *Any system of a finite number of convergent double sequences with the same limit is again convergent with the same limit, if by a subelement (superelement) of a system of double sequences be meant a common subelement (superelement) of these double sequences, and the notions of an interelement, lower limit, and limit of a system of double sequences be defined accordingly.*

In Theorems 36 and 37 the signs $\prod_{k=n}^{\infty} a_k$ and $\sum_{k=n}^{\infty} a_k$ are to be replaced by the signs $\prod\limits_{\substack{k,l=n\\(k,l)\,\epsilon\,\Theta}}^{\infty} a_{kl}$ and $\sum\limits_{\substack{k,l=n\\(k,l)\,\epsilon\,\Theta}}^{\infty} a_{kl}$ respectively. Also Theorems 39, 40, 41, and 42 are to be changed in a similar way.

A special case of Theorem 29′ is

THEOREM 123. *If a_{mn} $(m, n = 1, 2, 3, \cdots)$ are elements of a Boolean ring, m_k and n_k $(k = 1, 2, 3, \cdots)$ natural numbers such that $m_k \to \infty$ and $n_k \to \infty$ for $k \to \infty$, and* $\lim\limits_{m,n\to\infty} a_{mn}$ *exists, then* $\lim\limits_{k\to\infty} a_{m_k n_k}$ *exists too and is equal to* $\lim\limits_{m,n\to\infty} a_{mn}$.

It is an open question whether Theorem 123 can be inverted, i.e. whether $\lim\limits_{k\to\infty} a_{m_k n_k} = a$ whenever $m_k \to \infty$ and $n_k \to \infty$ for $k \to \infty$, implies $\lim\limits_{m,n\to\infty} a_{mn} = a$. We shall return to this point once more in §10.

9. Fundamental Sequences

DEFINITION 17. *A sequence of elements a_n of a Boolean ring A is said to be a fundamental sequence with respect to A if*

$$\underset{m,n\to\infty}{{}^{(A)}\lim}\,(a_m + a_n) = 0.$$

THEOREM 124. *A sequence a_n is a fundamental sequence if and only if the sequence $a_n b$ is a fundamental sequence for every element b.*

PROOF. See Definition 17, Theorem 5, and Lemma 5.

THEOREM 125. *A bounded sequence is a fundamental sequence if and only if the following condition is satisfied: if x is an element of the property that*

$$x < o + u$$

whenever u is a subelement, and o a superelement of the sequence, then $x = 0$.

PROOF. If u is a subelement and o a superelement of the sequence a_n then $o + u$ is a superelement of the double sequence $a_m + a_n$. Conversely to every superelement o^* of the double sequence $a_m + a_n$ there exist a subelement u and a superelement o of the sequence a_n such that

$$o^* = o + u. \tag{64}$$

For if o^* is a superelement of the double sequence $a_m + a_n$, there exists a natural number k such that

$$a_m + a_n < o^* \tag{65}$$

for $m \geqq k$, $n \geqq k$. From (65) follows

$$a_m + a_m o^* < a_n < a_m \vee o^*.$$

Hence $a_m + a_m o^*$ is a subelement, and $a_m \vee o^*$ a superelement of the sequence a_n, provided that $m \geqq k$. On the other side (64) is satisfied if we put $u = a_m + a_m o^*$, $o = a_m \vee o^*$.

After this is settled Theorem 125 follows easily from Theorem 22.

It is obvious that the fundamental sequences of a Boolean ring can be defined by their properties stated in Theorems 124 and 125. Compare Definitions 3 and 4.

THEOREM 126. *Any fundamental sequence has at the utmost one interelement.*

PROOF. If a and a^* are interelements of the sequence a_n then $a + aa^*$ and 0 are interelements of the bounded sequence $a_n(a + aa^*)$. Hence 0 is the only subelement of the latter sequence, and if o is a superelement of this sequence, then

$$a + aa^* < o.$$

The more we may say that

$$a + aa^* < o + u$$

whenever u is a subelement and o a superelement of the sequence $a_n(a + aa^*)$. If the sequence a_n is a fundamental sequence, the sequence $a_n(a + aa^*)$ is, by Theorem 124, likewise a fundamental sequence, and from Theorem 125 follows $a + aa^* = 0$. Just so we can prove that $a^* + aa^* = 0$. Hence $a^* = a$.

THEOREM 127. *Every convergent sequence is a fundamental sequence.*

PROOF. See Definition 17 and Relation (61) of Theorem 115.

On the other hand it is not true that every sequence without an interelement is a fundamental sequence. Let A be the Boolean ring of all finite classes of natural numbers, for $n = 1, 2, 3, \cdots$ $a_{2n-1} = \{1, 2, 3, \cdots, n\}$ (the class of the numbers $1, 2, 3, \cdots, n$), $a_2 = 0$ (the void class), and $(n = 2, 3, \cdots)$ $a_{2n} = \{2, 3, \cdots, n\}$. Then the sequence a_n $(n = 1, 2, 3, \cdots)$ has no interelement (there is no element a satisfying $a > u$ for all subelements u of the sequence), and it is not a fundamental sequence either.

THEOREM 128. *If $\mathfrak{a}$ is an invariant subring of a Boolean ring A, and a_n $(n = 1, 2, 3, \cdots)$ are elements of $\mathfrak{a}$, then the sequence a_n is a fundamental sequence with respect to $\mathfrak{a}$ if and only if it is a fundamental sequence with respect to A.*

PROOF. See Theorem 50 and Definition 17.

THEOREM 129. *A sequence of elements a_n of a Boolean ring A is a fundamental sequence with respect to A if and only if the sequence $\mathfrak{a}(a_n)$ is convergent with respect to $\mathfrak{N}$.*

PROOF. See Theorems 38, 68, 126, 127, and 128.

THEOREM 130. *Every monotonic sequence of elements of a Boolean ring is a fundamental sequence.*

PROOF. See Theorems 35 and 129.

DEFINITION 18. *A Boolean ring A is called σ-complete if every fundamental sequence in A converges with respect to A.*

THEOREM 131. *A Boolean ring A is σ-complete if and only if A is a Boolean σ-ring.*

PROOF. If all fundamental sequences of A converge, and a_n $(n = 1, 2, 3, \cdots)$ are elements of A, then from Theorems 32, 34, and 130 follows that $\prod_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} a_n$ exist. If A is a Boolean σ-ring, every fundamental sequence of A has, by Theorems 38 and 126, exactly one interelement.

D. van Dantzig (I) has shown how every topological ring, satisfying the neighbourhood-axioms of Fréchet-Hausdorff and having the property that sum, difference, and product of two elements of the ring are continuous functions of these elements, can be extended to a topological ring of the same kind and of the property that every fundamental sequence of the extended ring is convergent. In this paper we have defined for every Boolean ring A an intrinsic sequential topology $\tau(A)$, and have not considered any other sequential topology in A until now. Therefore if we consider an extension B of A we are interested only in the case that B is again a Boolean ring, and that the sequential topology $\tau(B)$ is an extension of the sequential topology $\tau(A)$. We know already that $\tau(B)$ is certainly an extension of $\tau(A)$ if B is an invariant extension of A. (Theorem 50.) And Theorems 69 and 131 show that any Boolean ring A can be invariantly extended to a Boolean ring B in which every fundamental sequence converges. Now we shall try to define the notion of a *minimal* invariant extension of A of this property, analogous to the notion of a minimal complete invariant extension of A, defined by Definition 12.

THEOREM 132. *If we mean by the sum of two sequences a_n and b_n in a Boolean ring A the sequence $a_n + b_n$, and by their product the sequence $a_n b_n$, the sequences*

in A form a Boolean ring S; the fundamental sequences in A form a subring F of S, the convergent sequences in A form a subring C of F, and the null sequences (sequences converging towards zero) *form an ideal* $\mathfrak{n}$ *in C.*

PROOF. The first proposition is evident. The second proposition follows from the equations

$$(a_m + b_m) + (a_n + b_n) = (a_m + a_n) + (b_m + b_n)$$

$$a_m b_m + a_n b_n = a_m(b_m + b_n) + (a_m + a_n)b_n$$

and from Theorems 92 and 115, the third from Theorem 115, and the fourth from Theorems 92 and 115.

Now let A_1 be a homomorphic image of F such that there exists a homomorphism γ from F to A_1 determining exactly the ideal $\mathfrak{n}$ of F. (I.e. $\mathfrak{n}$ shall be the class of the elements of F carried by γ into the zero of A_1.) Such a homomorphic image of F is for instance the quotient-ring $F/\mathfrak{n}$. Two elements of C are carried by γ into the same element of A_1 if and only if they are sequences convergent with respect to A and having the same limit with respect to A. Hence if we denote the image of C in A_1 by C^* and assign to every element a of A that element of C^* in which the convergent sequence, having all terms equal to a, is carried by γ, then this correspondence is an isomorphism between A and C^*. Hence A_1 can be chosen in such a way that the image of the sequence $\{a, a, a, \cdots\}$ in A_1 is identical with a itself. In this case A_1 is an extension of A.

DEFINITION 19. *Let the signs A, F, C, and* $\mathfrak{n}$ *have the same meaning as in Theorem* 132, *and let* A_1 *be a homomorphic image of F, and* γ *a homomorphism* $F \to A_1$ *determining exactly the ideal* $\mathfrak{n}$ *of F and having the property that if a is an element of A, the element of* A_1 *in which* γ *carries the sequence having all terms equal to a, is identical with a. In this case we call* A_1 *a first fundamental extension of A.*

Two fundamental sequences a_n and b_n have the same image in A_1 if and only if their difference (this is here the same as $a_n + b_n$) is a null sequence. D. van Dantzig calls such fundamental sequences *concurrent.*

LEMMA 10. *If* A_1 *is a first fundamental extension of a Boolean ring A, then A is invariant in* A_1.

PROOF. Let $\mathfrak{B}$ be a subclass of A, and b an element of A such that

$$\sum_{a \in \mathfrak{B}}^{(A)} a = b.$$

Further let f be an element of A_1 such that

$$f > a$$

for all elements a of $\mathfrak{B}$. If a_n is a fundamental sequence in A, of which the image in A_1 is f, then the sequence $a_n a + a$ must be contained in $\mathfrak{n}$, i.e. we must have

$$\lim_{n \to \infty}^{(A)} a_n a = a.$$

Hence we have, by Theorem 41,

$$\sum_{k=n}^{\infty}{}^{(A)}\, a_k a = a$$

and

$$\sum_{a \in \mathfrak{B}}{}^{(A)} \sum_{k=n}^{\infty}{}^{(A)}\, a_k a = b$$

for $n = 1, 2, 3, \cdots$. Since the joins $\sum_{a \in \mathfrak{B}}{}^{(A)}\, a_k a$ exist (see Theorem 13) we can easily conclude that also

$$\sum_{k=n}^{\infty}{}^{(A)} \sum_{a \in \mathfrak{B}}{}^{(A)}\, a_k a = b.$$

Hence we have

$$\sum_{k=n}^{\infty}{}^{(A)}\, a_k b = b$$

and, by Theorem 41,

$${}^{(A)}\overline{\lim_{n \to \infty}}\, a_n b = b,$$

whence

$${}^{(A)}\lim_{n \to \infty}\, a_n b = b$$

because $a_n b$ is a fundamental sequence. Thus we have $fb = b$ or $f > b$. This proves that

$$\sum_{a \in \mathfrak{B}}{}^{(A_1)}\, a = b.$$

THEOREM 133. *If the signs A, F, γ, and A_1 have the same meaning as in Definition 19 then any fundamental sequence in A converges with respect to A_1 to that element of A_1 which is its image by virtue of γ.*

PROOF. First we consider a monotonic sequence in A. Let $a_n (n = 1, 2, 3, \cdots)$ be elements of A, and for instance $a_k < a_l$ for $k < l$, $k, l = 1, 2, 3, \cdots$. We have

$${}^{(A)}\lim_{n \to \infty}\, a_n a_k = a_k . \qquad (k = 1, 2, 3, \cdots .)$$

Hence we have, if f is the image of the fundamental sequence a_n, for $k = 1, 2, 3, \cdots$

$$f a_k = a_k$$

or

$$f > a_k .$$

If g is an element of A_1, b_n a fundamental sequence in A the image of which is g, and $g > a_k$ for $k = 1, 2, 3, \cdots$, we have for $k = 1, 2, 3, \cdots$

$$^{(A)}\lim_{n\to\infty}(a_k + a_k b_n) = 0,$$

whence

$$\mathfrak{a}(a_k) + \mathfrak{a}(a_k)\,{}^{(\mathfrak{R})}\lim_{n\to\infty}\mathfrak{a}(b_n) = 0,$$

$$^{(\mathfrak{R})}\lim_{n\to\infty}\mathfrak{a}(a_n) + {}^{(\mathfrak{R})}\lim_{n\to\infty}\mathfrak{a}(a_n).\,{}^{(\mathfrak{R})}\lim_{n\to\infty}\mathfrak{a}(b_n) = 0,$$

$$^{(A)}\lim_{n\to\infty}(a_n + a_n b_n) = 0,$$

and

$$g > f.$$

This proves

$$\sum_{k=1}^{\infty}{}^{(A_1)}\, a_k = {}^{(A_1)}\lim_{n\to\infty} a_n = f.$$

If $a_k > a_l$ for $k < l$, $k, l = 1, 2, 3, \cdots$, and f is again the image of the fundamental sequence a_n, we can prove in a similar way that

$$\prod_{k=1}^{\infty}{}^{(A_1)}a_k = {}^{(A_1)}\lim_{n\to\infty} a_n = f.$$

Now let a_n be an arbitrary fundamental sequence in A, and f its image. From what we have just proved follows that $\prod_{l=1}^{\infty}{}^{(A_1)}\, a_{k+l-1}$ and $\sum_{l=1}^{\infty}{}^{(A_1)}\, a_{k+l-1}$ exist and are the images of the fundamental sequences $\prod_{l=1}^{n} a_{k+l-1}$ $(n = 1, 2, 3, \cdots)$ and $\sum_{l=1}^{n} a_{k+l-1}$ $(n = 1, 2, 3, \cdots)$ respectively. But we have

$$^{(A)}\lim_{n\to\infty}\left(a_n \prod_{l=1}^{n} a_{k+l-1} + \prod_{l=1}^{n} a_{k+l-1}\right) = 0$$

and

$$^{(A)}\lim_{n\to\infty}\left(a_n + a_n \sum_{l=1}^{n} a_{k+l-1}\right) = 0.$$

Hence

$$\prod_{l=1}^{\infty}{}^{(A_1)}\, a_{k+l-1} < f < \sum_{l=1}^{\infty}{}^{(A_1)}\, a_{k+l-1}.$$

From this relation it can be easily concluded that f is an interelement of the sequence a_n with respect to A_1. Since this sequence is, by Theorem 128 and Lemma 10, also with respect to A_1 a fundamental sequence, Theorem 133 is proved. (See Theorem 126.)

LEMMA 11. *If A_1 is a first fundamental extension of a Boolean ring A, then A_1 is also a join-extension of A.*

PROOF. A_1 is, by Lemma 10, an invariant extension of A. Let N be the maximal join-extension of A invariant in A_1. Then we have, by Theorem 121,

$$N \supset A_1$$

because all elements of A_1 are, by Theorem 133, limits of elements of A and by this limits of elements of N with respect to A_1. Hence $N = A_1$.

In the paper (I) the ring which corresponds to the ring denoted here by A_1, has already the property that all its fundamental sequences converge. We shall see that in our case this assertion is in general not true though D. van Dantzig's "Ringkomplettierungsaxiom" (the product of a null sequence and a fundamental sequence is a null sequence) is satisfied, according to Theorem 92.

LEMMA 12. *If A and A^* are Boolean rings between which there exists an isomorphism Λ, and A_1 and A_1^* are first fundamental extensions of A and A^* respectively, then A_1 and A_1^* are isomorphic too, and the isomorphism Λ can be extended to an isomorphism Λ_1 between A_1 and A_1^* in exactly one way.*

PROOF. Λ_1 can be defined in the following way: let f be an element of A_1, a_n a sequence of elements of A converging towards f with respect to A_1, a_n^* the sequence in A^* corresponding to the sequence a_n by virtue of Λ, and

$$f^* = {}^{(A_1^*)}\lim_{n\to\infty} a_n^*;$$

this correspondence $f \to f^*$ is the desired isomorphism Λ_1.

In the following we shall denote by ρ the smallest ordinal number which is the ordinal type of an uncountable set. More precisely let ψ be the ordinal number of some uncountable well-ordered set, and ρ the smallest ordinal number satisfying $\rho \leqq \psi$ which is as well the ordinal type of an uncountable set.

DEFINITION 20. *Let A be an arbitrary Boolean ring, and α an ordinal number satisfying $\alpha \leqq \rho$. Let us assign a Boolean ring A_μ to any ordinal number μ satisfying $\mu \leqq \alpha$ by the following transfinite induction:*

1.) *A_1 is a first fundamental extension of A;*

2.) *if the ordinal number μ has a predecessor ν (thus $\mu = \nu + 1$) then A_μ is a first fundamental extension of A_ν;*

3.) *if the ordinal number μ is a limit-number then*

$$A_\mu = \sum_{\nu<\mu} A_\nu,$$

the sign Σ denoting the set-theoretical union; the ring-operations in A_μ are defined in the following way: if a and b are elements of A_μ, and ν an ordinal number such that $\nu < \mu$, $a \in A_\nu$, $b \in A_\nu$, then

$$a \Delta_{A_\mu} b = a \Delta_{A_\nu} b,$$

$$a \wedge_{A_\mu} b = a \wedge_{A_\nu} b.$$

In this case we call A_α a fundamental extension of the Boolean ring A of the order α.

If μ is an ordinal number satisfying $\mu < \alpha$, the Boolean ring A_μ appearing in Definition 20 is obviously a fundamental extension of A of the order μ.

THEOREM 134. *If A is an arbitrary Boolean ring, α an ordinal number satisfying $\alpha \leqq \rho$, and A_α a fundamental extension of A of the order α, then A_α is a join-extension of A.*

PROOF. See Theorems 62 and 63, Lemma 11, and Definition 20.

THEOREM 135. *If A is an arbitrary Boolean ring, α and β are ordinal numbers satisfying $\beta < \alpha \leqq \rho$, A_α and A_β fundamental extensions of A of the orders α and β respectively, and $A_\beta \subset A_\alpha$, then A_α is a join-extension of A_β.*

PROOF. See Theorems 64 and 134.

LEMMA 13. *Let A be a subring of a Boolean ring B, and $A_{(1)}$ the class of all elements of B which are limits of sequences of elements of A with respect to B. Then $A_{(1)}$ is a subring of B containing A. If A is especially invariant in B then $A_{(1)}$ is a join-extension of A invariant in B.*

PROOF. It is evident that $A_{(1)} \supset A$. That $A_{(1)}$ is a subring of B follows from Theorem 115. If A is invariant in B, Theorem 121 involves that $A_{(1)}$ is contained in the maximal join-extension of A invariant in B.

DEFINITION 21. *If the signs A, B, and $A_{(1)}$ have the same meaning as in Lemma 13 then $A_{(1)}$ is called the first limit-ring of A in B.*

DEFINITION 22. *Let A be a subring of a Boolean ring B, and α an ordinal number satisfying $\alpha \leqq \rho$. Let us assign a subring $A_{(\mu)}$ of B to any ordinal number μ satisfying $\mu \leqq \alpha$ by the following transfinite induction:*

1.) *$A_{(1)}$ is the first limit-ring of A in B;*

2.) *if the ordinal number μ has a predecessor ν (so that $\mu = \nu + 1$) then $A_{(\mu)}$ is the first limit-ring of $A_{(\nu)}$ in F;*

3.) *if the ordinal number μ is a limit-number then*

$$A_{(\mu)} = \sum_{\nu<\mu}^{(\mathfrak{R})} A_{(\nu)},$$

by $\mathfrak{R}$ denoted the lattice of all subrings of B.

In this case we call $A_{(\alpha)}$ the limit-ring of A in B of the order α.

THEOREM 136. *If A is an invariant subring of a Boolean ring B, then any limit-ring $A_{(\alpha)}$ of A in B is invariant in B and a join-extension of A and of every limit-ring of A in B of lower order.*

PROOF. From Definition 22 it can be easily concluded that $A_{(\alpha)}$ is contained in the maximal join-extension of A invariant in B. For the rest see Theorems 53, 61, 64, and 70.

THEOREM 137. *If A is any Boolean ring, α an ordinal number satisfying $\alpha \leqq \rho$, A_α a fundamental extension of A of the order α, and B an invariant extension of A_α, then A_α is the limit-ring of A in B of the order α.*

PROOF. Assign a Boolean ring A_μ to any ordinal number μ satisfying $\mu < \alpha$ in such a way that the rings A_μ ($\mu \leqq \alpha$) satisfy the conditions stated in Definition 20. Let us suppose that $\mu_0 < \alpha$, and that we have already proved that A_{μ_0} is

the limit-ring of A in B of the order μ_0. Then let a_n be a sequence of elements of A_{μ_0} which has a limit a with respect to B. Since A_{μ_0} is, by Theorems 53, 61, and 135, invariant in B, the sequence a_n is a fundamental sequence with respect to A_{μ_0} according to Theorems 127 and 128. Hence this sequence has, by Definitions 19 and 20 and by Theorem 133, a limit with respect to A_{μ_0+1}, and this limit is necessarily identical with a, because also A_{μ_0+1} is invariant in B. Hence $a \in A_{\mu_0+1}$. If we assume conversely that a is an element of A_{μ_0+1} then a is, by Definitions 19 and 20 and Theorem 133, the limit of some sequence of elements of A_{μ_0} with respect to A_{μ_0+1}, and consequently also with respect to B because A_{μ_0+1} is invariant in B. Thus we find that a is an element of the limit-ring of A in B of the order $\mu_0 + 1$. Hence this limit-ring is identical with A_{μ_0+1}. We see that Theorem 137 can be proved by transfinite induction; for the conclusion from the predecessors of a limit-number to this limit-number itself is still simpler.

THEOREM 138. *If A is any Boolean ring, α and β ordinal numbers satisfying $\beta \leqq \alpha \leqq \rho$, and A_α a fundamental extension of A of the order α, then there exists exactly one subring A_β of A_α which is a fundamental extension of A of the order β.*

PROOF. The existence of A_β follows from Definition 20, the uniqueness from Theorem 137. (A_β is the limit-ring of A in A_α of the order β.)

THEOREM 139. *If A is an invariant subring of a Boolean σ-ring B, α an ordinal number satisfying $\alpha \leqq \rho$, and $A_{(\alpha)}$ the limit-ring of A in B of the order α, then $A_{(\alpha)}$ is a fundamental extension of A of the order α.*

PROOF. If a_n ($n = 1, 2, 3, \cdots$) are elements of A, and the sequence consisting of these elements is a fundamental sequence with respect to A, it is, by Theorem 128, also a fundamental sequence with respect to B; hence it is, by Theorem 131, convergent with respect to B; its limit is an element of $A_{(1)}$. Conversely any element of $A_{(1)}$ can be obtained in this way. If we assign to every fundamental sequence in A that element of $A_{(1)}$ which is its limit with respect to B we get a homomorphism γ from the Boolean ring F of the fundamental sequences in A to $A_{(1)}$ of the sort demanded in Definition 19. This proves the theorem for the case $\alpha = 1$. The general theorem follows now by transfinite induction.

THEOREM 140. *Let A and A^* be Boolean rings between which there exists an isomorphism Λ, α and β ordinal numbers satisfying $\beta \leqq \alpha \leqq \rho$, A_β a fundamental extension of A of the order β, and A^*_α a fundamental extension of A^* of the order α. Then there exists exactly one subring R^* of A^*_α such that the isomorphism Λ can be extended to an isomorphism Λ_β between A_β and R^*. Λ_β is uniquely determined, and R^* is a fundamental extension of A^* of order β. If especially $\beta = \alpha$ then $R^* = A^*_\alpha$.*

PROOF. Let A^*_β be a subring of A^*_α which is a fundamental extension of A^* of the order β. That A^*_β exists and is unique follows from Theorem 138. If we put $R^* = A^*_\beta$ it ensues from Lemma 12 and by transfinite induction that the isomorphism Λ_β, demanded by Theorem 140, exists and is unique. The uniqueness of R^* is involved by Theorem 83.

THEOREM 141. *If A is a subring of a Boolean ring B the closure of A in B is as well a subring of B; it is identical with the limit-ring $A_{(\rho)}$ of A in B of the order ρ.*

PROOF. Let a_n be a sequence of elements of $A_{(\rho)}$ which has a limit a with re-

spect to B. Further, for $n = 1, 2, 3, \cdots$ let α_n be the smallest ordinal number which is not larger than ρ and has the property that a_n is contained in $A_{(\alpha_n)}$, by $A_{(\beta)}$ $(\beta \leqq \rho)$ denoted the limit-ring of A in B of the order β. Since ρ is a limit-number we have certainly

$$\alpha_n \neq \rho$$

for $n = 1, 2, 3, \cdots$. Further let α be the smallest ordinal number satisfying the inequality

$$\alpha_n \leqq \alpha \leqq \rho$$

for $n = 1, 2, 3, \cdots$. We have

$$\alpha \neq \rho;$$

for the asumption $\alpha = \rho$ would imply that ρ would be the ordinal number of a finite or countable set, contrary to the definition of ρ. Hence we have also

$$\alpha + 1 \leqq \rho.$$

Now the elements a_n $(n = 1, 2, 3, \cdots)$ are all contained in $A_{(\alpha)}$. Consequently the element

$$a = {}^{(B)}\lim_{n \to \infty} a_n$$

is contained in $A_{(\alpha+1)}$. Thus we have

$$a \in A_{(\rho)}$$

and it is proved that $A_{(\rho)}$ is a closed subclass of B. On the other side it can be shown by transfinite induction that any subclass of B closed under $\Gamma(\tau(B))$ and containing A contains $A_{(\alpha)}$ whenever $\alpha \leqq \rho$, and by this especially $A_{(\rho)}$. Thus Theorem 141 is proved.

If B is especially the Boolean ring of all sets of real numbers, and A the subclass of B defined by the property that a is an element of A if and only if it is either the empty set or the union of a finite number of sets, every one of which is either an open real interval or consists of exactly one real number (see the example stated after Definition 9), then $A_{(\rho)}$ is the Boolean ring of the Borel sets of real numbers. In this case it is known that the rings $A_{(\alpha)}$ $(\alpha \leqq \rho)$ are all different. Since B is a complete Boolean ring here, $A_{(\alpha)}$ is, by Theorem 139, a fundamental extension of A of the order α. It follows that if A_1 is a first fundamental extension of A, A_1 is not a σ-complete Boolean ring. If we compare this result with the results of §7 of the paper (I), we find that the sequential topology $\tau(A)$ of a Boolean ring A defined by Definition 8, in general cannot be derivated from a neighbourhood-topology satisfying the axioms of Fréchet-Hausdorff.

THEOREM 142. *If A_ρ is a fundamental extension of a Boolean ring A of the order ρ then A_ρ is a Boolean σ-ring.*

PROOF. Let B be a Boolean σ-ring and an invariant extension of A_ρ. That such a Boolean ring B exists follows from Theorem 69. Then A_ρ is, by Theorem 137, the limit-ring of A in B of the order ρ, and it follows from Theorem 141 that A_ρ is closed in B. Hence A_ρ is a Boolean σ-ring, according to Theorem 117.

DEFINITION 23. *An extension $\tilde{A}$ of a Boolean ring A is said to be a minimal invariant σ-extension of A, or shorter: a σ-completion of A, if $\tilde{A}$ is a Boolean σ-ring invariant over A, and no proper subring of $\tilde{A}$ containing A is itself a Boolean σ-ring.*

THEOREM 143. *An extension $\tilde{A}$ of a Boolean ring A is a minimal invariant σ-extension of A if and only if $\tilde{A}$ is a fundamental extension of A of the order ρ.*

PROOF. If $\tilde{A}$ is a minimal invariant σ-extension of A, the limit-ring of A in $\tilde{A}$ of the order ρ is, by Theorem 139, a fundamental extension of A of the order ρ; hence it is, by Theorem 142, a Boolean σ-ring, and Definition 23 involves that it cannot be different from $\tilde{A}$. If $\tilde{A}$ is a fundamental extension of A of the order ρ then $\tilde{A}$ is, by Theorem 142, a Boolean σ-ring and, by Theorem 134, a join-extension and by this an invariant extension of A. (Theorem 61.) On the other side $\tilde{A}$ is, by Theorems 137 and 141, the closure of A in $\tilde{A}$. If R is a σ-subring of $\tilde{A}$ containing A then R is, by Theorems 61 and 64, invariant in $\tilde{A}$ and, by Theorem 118, closed in $\tilde{A}$. Hence $R = \tilde{A}$, and it is proved that $\tilde{A}$ is a σ-completion of A.

THEOREM 144. *If $\tilde{A}$ is a σ-completion of a Boolean ring A, and B an invariant extension of $\tilde{A}$, then $\tilde{A}$ is the closure of A in B.*

PROOF. See Theorems 137, 141, and 143.

THEOREM 145. *If A is an invariant subring of a Boolean σ-ring B, the closure of A in B is a minimal invariant σ-extension of A.*

PROOF. See Theorems 139, 141, and 143.

THEOREM 146. *If $\tilde{A}$ and $\tilde{A}^*$ are minimal invariant σ-extensions of the same Boolean ring A then $\tilde{A}$ and $\tilde{A}^*$ are isomorphic, and there is exactly one isomorphism between $\tilde{A}$ and $\tilde{A}^*$ carrying every element of A into itself.*

Hence we may say that $\tilde{A}$ and $\tilde{A}^*$ are abstractly identical.

PROOF. See Theorems 140 and 143.

10. The Sequential Topology $\tau(\Gamma(\tau(A)))$

If A is any Boolean ring, the closure-topology $\Gamma(\tau(A))$ in A, considered in §7, determines a new sequential topology $\tau(\Gamma(\tau(A)))$ in A according to (III), Theorem 15. The closure-topology determined by this sequential topology is identical with $\Gamma(\tau(A))$ or, in other words, the closure-topology $\Gamma(\tau(A))$ and the sequential topology $\tau(\Gamma(\tau(A)))$ are equivalent. (See (III), Definition 8 and Theorem 18.)

LEMMA 14. *If a sequence a_n has a limit a under $\tau(A)$ it has also under $\tau(\Gamma(\tau(A)))$ the limit a and no other limit.*

PROOF. It is obvious that the sequence a_n has the limit a also under $\tau(\Gamma(\tau(A)))$. If $b \neq a$, at the utmost a finite number of terms of the sequence a_n is equal to b, and Theorem 29 involves that the terms of this sequence different from b and the element a together constitute a set closed under $\Gamma(\tau(A))$. The

complement of this set is an open set which contains b, but fails to contain infinitely many terms of the sequence a_n. Hence b is certainly no limit of the sequence a_n under $\tau(\Gamma(\tau(A)))$. This completes the proof of Lemma 14.

Further it is obvious that our Theorems 25, 29, and 30 hold also for the sequential topology $\tau(\Gamma(\tau(A)))$. (They hold in any derivative sequential topology.)

THEOREM 147. (Compare (VII), Satz 29.) *If a_n $(n = 1, 2, 3, \cdots)$ and a are elements of a Boolean ring A then a is a limit of the sequence a_n under $\tau(\Gamma(\tau(A)))$ if and only if every partial sequence of the sequence a_n contains a further partial sequence converging towards a under $\tau(A)$. Hence any sequence of elements of A has at the utmost one limit under $\tau(\Gamma(\tau(A)))$.*

PROOF. If a set open under $\Gamma(\tau(A))$ and containing a fails to contain an infinite partial sequence of the given sequence, no partial sequence of the mentioned partial sequence can converge towards a under $\tau(A)$. Hence if every partial sequence of the given sequence contains a further partial sequence converging towards a under $\tau(A)$, the given sequence has certainly the limit a under $\tau(\Gamma(\tau(A)))$.

Let us suppose conversely that the sequence a_n has the limit a under $\tau(\Gamma(\tau(A)))$. If the sequence a_n had not the property asserted in Theorem 147 we could state a partial sequence a_{n_k} $(k = 1, 2, 3, \cdots)$ of this sequence containing no further partial sequence converging towards a under $\tau(A)$. Then the generalized Theorem 29 and Lemma 14 involve that the sequence a_{n_k} contains no sequence convergent under $\tau(A)$ at all. Moreover we may suppose that the elements a_{n_k} are all different from a. Thus the elements of A which are different from all elements a_{n_k} form an open set which contains a, but fails to contain infinitely many terms of the sequence a_n. This result would contradict to the supposition that the sequence a_n converges towards a under $\tau(\Gamma(\tau(A)))$.

DEFINITION 24. *If a_n $(n = 1, 2, 3, \cdots)$ and a are elements of a Boolean ring A, we say that the sequence a_n $(n = 1, 2, 3, \cdots)$ converges weakly towards a with respect to A, or that a is its weak limit with respect to A, and write*

$$^{(A)}\operatorname*{Lim}_{n\to\infty} a_n = a$$

if the sequence a_n converges towards a under $\tau(\Gamma(\tau(A)))$.

The words "with respect to A" and the superscript in the sign $^{(A)}\operatorname*{Lim}_{n\to\infty} a_n$ may be omitted if there is no doubt which Boolean ring is meant.

We have chosen the words "weak convergence" and "weak limit" for convenience although the considered convergence under $\tau(\Gamma(\tau(A)))$ is no analogue of the weak convergence in a linear metric space.

As far as the sequential topology $\tau(A)$ is concerned, we shall use the same terminology and denotation in this paragraph as in the preceding ones.

The sequential topology $\tau(\Gamma(\tau(A)))$ is in general different from the sequential topology $\tau(A)$. Let A be the Boolean ring of those sets of real numbers which

are either empty or unions of finite numbers of right-hand open intervals. (A is identical with the Boolean ring $\mathfrak{b}$ considered after Definition 9.) If μ and ν are real numbers we denote the interval $\mu \leqq \xi < \nu$ by $< \mu, \nu)$. Now put

$$a_n = < \frac{n - 2^m}{2^m}, \frac{n - 2^m + 1}{2^m})$$

for $2^m \leqq n < 2^{m+1}$, $m = 0, 1, 2, 3, \cdots$. a_n $(n = 1, 2, 3, \cdots)$ are elements of A. If m and n are natural numbers, at least one of the equations

$$a_m < a_n ,$$

$$a_m > a_n ,$$

and

$$a_m a_n = 0$$

is valid. To a given natural number n_0 there exist at the utmost finitely many natural numbers n such that

$$a_n > a_{n_0} .$$

Hence any partial sequence b_n of the sequence a_n contains at least either a partial sequence c_n satisfying

$$c_{n+1} < c_n \qquad (n = 1, 2, 3, \cdots) \tag{67}$$

or a partial sequence c_n satisfying

$$c_k c_l = 0 \qquad (k, l = 1, 2, 3, \cdots, k \neq l). \tag{68}$$

(67) as well as (68) implies

$$\lim_{n \to \infty} c_n = 0.$$

(See Theorem 116.) Hence we have

$$\operatorname*{Lim}_{n \to \infty} a_n = 0.$$

On the other side we have

$$\sum_{n=2^m}^{2^{m+1}-1} a_n = a_1 \qquad (m = 1, 2, 3, \cdots),$$

thus

$$\sum_{k=n}^{\infty} a_k = a_1 \qquad (n = 1, 2, 3, \cdots)$$

and

$$\overline{\lim_{n \to \infty}} a_n = a_1 .$$

Since $a_1 \neq 0$ the equation $\lim_{n \to \infty} a_n = 0$ is not valid.

THEOREM 148. *If* $\operatorname{Lim}_{n\to\infty} a_n$ *exists then it is an interelement of the sequence* a_n *and of any partial sequence of this sequence.*

PROOF. See Theorems 27 and 147.

Now it is obvious that our Theorems 50, 87, 92, and 115 are valid also for the weak convergence. The same holds for Theorem 91 if it is pronounced only for convergent sequences. I.e. we have

THEOREM 149. *If* $a_n < b_n$ *for* $n = 1, 2, 3, \cdots$, *and* $\operatorname{Lim}_{n\to\infty} a_n$ and $\operatorname{Lim}_{n\to\infty} b_n$ *exist, then*

$$\operatorname{Lim}_{n\to\infty} a_n < \operatorname{Lim}_{n\to\infty} b_n .$$

PROOF. We have for $n = 1, 2, 3, \cdots$

$$a_n = a_n b_n ,$$

hence, by the generalized Theorem 115,

$$\operatorname{Lim}_{n\to\infty} a_n = \operatorname{Lim}_{n\to\infty} a_n \cdot \operatorname{Lim}_{n\to\infty} b_n$$

proving Theorem 149.

If a_{mn} is a double sequence in a Boolean ring A, Θ the class of ordered couples of natural numbers for which the sign a_{mn} is defined, and a an element of A, we say that the double sequence a_{mn} converges weakly towards a, or has the weak limit a, and write

$$^{(A)}\operatorname{Lim}_{\substack{m,n\to\infty \\ (m,n)\,\epsilon\,\Theta}} a_{mn} = a,$$

if to any open set containing a there can be assigned a natural number k such that a_{mn} is contained in this open set for $m \geqq k$, $n \geqq k$, $(m, n) \epsilon \Theta$. If Θ is the class of all ordered couples of natural numbers, we write simply $^{(A)}\operatorname{Lim}_{m,n\to\infty} a_{mn}$ instead of $^{(A)}\operatorname{Lim}_{\substack{m,n\to\infty \\ (m,n)\,\epsilon\,\Theta}} a_{mn}$. In both cases the superscript (A) may be omitted if there is no doubt which Boolean ring is meant.

Theorems 25, 29′ (8), and 123 hold also for weakly-convergent double sequences. Moreover Theorem 123 pronounced for weakly-convergent double sequences can be inverted. I.e. if $m_k \to \infty$, $n_k \to \infty$ for $k \to \infty$ implies $\operatorname{Lim}_{k\to\infty} a_{m_k n_k} = a$, then $\operatorname{Lim}_{m,n\to\infty} a_{mn} = a$. (To prove this proposition is not difficult.) In particular if $m_k \to \infty$, $n_k \to \infty$ for $k \to \infty$ implies $\lim_{k\to\infty} a_{m_k n_k} = a$, we have certainly $\operatorname{Lim}_{m,n\to\infty} a_{mn} = a$. On the other hand compare the remark after Theorem 123.—Any double sequence of a Boolean ring A has at the utmost one weak limit with respect to A.

Now it is clear that also Theorems 50, 87, 92, 115 and 149 hold for weakly-convergent double sequences.

DEFINITION 25. *A sequence of elements a_n of a Boolean ring A is said to be a weak fundamental sequence with respect to A if*

$$\overset{(A)}{\operatorname{Lim}_{m,n\to\infty}}(a_m + a_n) = 0.$$

Theorems 127, 128, and 132 continue to be right if we replace the notions of "fundamental sequence" and "convergence" wherever they appear, by the notions of "weak fundamental sequence" and "weak convergence."

THEOREM 150. *Any two weakly-concurrent fundamental sequences are concurrent.*

PROOF. If a_n and b_n are fundamental sequences, and

$$\operatorname{Lim}_{n\to\infty}(a_n + b_n) = 0,$$

then there exists, by Theorem 147, a sequence n_k of natural numbers different from each other such that

$$\lim_{k\to\infty}(a_{n_k} + b_{n_k}) = 0. \tag{69}$$

On the other side we have, by Theorem 123 and Definition 17,

$$\lim_{k\to\infty}(a_k + a_{n_k}) = 0 \tag{70}$$

and

$$\lim_{k\to\infty}(b_k + b_{n_k}) = 0. \tag{71}$$

From (69), (70), and (71) we get, according to Theorem 115,

$$\lim_{k\to\infty}(a_k + b_k) = 0.$$

THEOREM 151. *If b_n is a fundamental sequence, and a_n a sequence of the property that every partial sequence of the sequence a_n contains a fundamental sequence weakly-concurrent with the fundamental sequence b_n, then a_n is a weak fundamental sequence weakly-concurrent with the fundamental sequence b_n.*

PROOF. First we observe that any partial sequence of a fundamental sequence is a fundamental sequence concurrent with the whole sequence. This can be easily concluded from Definition 17 and from Theorems 29′ and 123. Now if we notice Theorems 147 and 150, we find that the suppositions of Theorem 151 involve

$$\operatorname{Lim}_{n\to\infty}(a_n + b_n) = 0.$$

Since

$$a_m + a_n = (a_m + b_m) + (a_n + b_n) + (b_m + b_n),$$

and

$$\lim_{m,n\to\infty}(b_m + b_n) = 0,$$

we have

$$\lim_{m,n\to\infty}(a_m + a_n) = 0,$$

according to the generalized Theorem 115.

On the other hand it is an open question whether Theorem 151 can be inverted, i.e. whether every weak fundamental sequence of a Boolean ring contains a fundamental sequence in the sense of Definition 17. If this inversion is true, it follows that every weak fundamental sequence of a Boolean σ-ring is weakly-convergent. If it were not true then it could occur that a Boolean ring (if you want even a complete one) could not be invariantly extended to a Boolean ring in which every weak fundamental sequence is weakly-convergent.

PRAGUE

REFERENCES

(I) D. van Dantzig: *Zur topologischen Algebra.* I. Komplettierungstheorie. Mathematische Annalen 107(1933), pp. 587–626.

(II) Garrett Birkhoff: *On the Structure of Abstract Algebras.* Proceedings of the Cambridge Philosophical Society 31(1935), pp. 433–454.

(III) Garrett Birkhoff: *On the Combination of Topologies.* Fundamenta Mathematicae 26(1936), pp. 156–166.

(IV) Fritz Klein: *Boole-Schrödersche Verbände.* Deutsche Mathematik 1 (1936), pp. 528–537.

(V) M. H. Stone: *The Theory of Representations for Boolean Algebras.* Transactions of the American Mathematical Society 40(1936), pp. 37–111.

(VI) Garrett Birkhoff: *Rings of Sets.* Duke Mathematical Journal 3(1937), pp. 443–454.

(VII) L. V. Kantorovitch: *Lineare halbgeordnete Räume.* Recueil Mathématique 2/44(1937), pp. 121–168.

(VIII) Garrett Birkhoff: *Lattices and their Applications.* Bulletin of the American Mathematical Society 44(1938), pp. 793–800.

ANNALS OF MATHEMATICS
Vol. 42, No. 5, December, 1941

ON INCIDENCE MATRICES, NUCLEI AND HOMOTOPY TYPES

BY J. H. C. WHITEHEAD

(Received January 16, 1940)

1. The main object of this paper is to establish as close a connection as possible between certain purely algebraic processes on the one hand, and the theory of nuclei[1] and homotopy types[2] of complexes on the other. The algebraic processes include the transformations, operating on the incidence matrices with elements in the group ring of a given complex, by means of which Reidemeister's invariants are defined.[3] But in the main theorems, namely theorems 8 and 9 in §4, the incidence matrices $\mathbf{r}^n, \cdots, \mathbf{r}^1$ are replaced by what we call a natural system for a given complex $\mathbf{K}^n$. A natural system, $(\mathbf{r}, R)$, consists of the matrices $\mathbf{r}^n, \cdots, \mathbf{r}^3$, together with a "natural" system, R, of generators and relations, for the fundamental group $\pi_1(\mathbf{K}^n)$. Such systems, defined purely algebraically, are classified by two kinds of equivalence, called L-equivalence and L^*-equivalence, of which the former implies the latter. Theorem 8 asserts two kinds of combinatorial invariance. It states that natural systems for two given complexes $\mathbf{K}_0^n$ and $\mathbf{K}_1^n$ are "L-equivalent" if $\mathbf{K}_0^n$ and $\mathbf{K}_1^n$ have the same nucleus, and "L^*-equivalent" if $\mathbf{K}_0^n$ and $\mathbf{K}_1^n$ are of the same homotopy type. Theorem 9 is complementary to theorem 8, and states that the elementary transformations by means of which X-equivalence is defined ($X = L$ or L^*), can be "copied geometrically." More precisely, if $(\mathbf{r}, R)$ is a natural system for $\mathbf{K}^n$, then any system which is X-equivalent to $(\mathbf{r}, R)$, is a natural system for some complex of the same homotopy type as $\mathbf{K}^n$, and for one which has the same nucleus as $\mathbf{K}^n$ if $X = L$. This means that one can carry out certain types of algebraic calculation without destroying the geometrical significance of the system on which one is operating, and the result is a step towards translating the theory of nuclei and m-groups, likewise homotopy types (of polyhedra), into purely algebraic terms.

As an application of the theorems in §§3 and 4 we prove, in §5, that two lens

---

[1] See J. H. C. Whitehead, Proc. L. M. S., 45 (1939), 243-327, which will be referred to as S.S. Throughout this paper we shall only be interested in the nucleus and homotopy type of a complex as principles of classification.

[2] See pp. 124-5 of W. Hurewicz, Proc. Kon. Akad. Amsterdam, 39 (1936), 117-25, or note III_1 in §6 below. Any two complexes with the same nucleus are of the same homotopy type.

[3] See, among other papers, K. Reidemeister, Abh. Hamb. Sem., 10 (1934), 211-15; 11 (1935), 102-9 and Journal für die r. u. a. Math., 173 (1935), 164-73. By the group ring of a complex we mean the group ring of its fundamental group, with (rational) integral coefficients. We confine ourselves to the incidence matrices which are defined in terms of the universal covering complex, leaving aside the concept of "Ueberdeckungen."

spaces of types[4] (m, q) and (m, r) are of the same homotopy type if they have the same intersection invariant,[5] that is to say if $r \equiv \pm l^2 q \pmod{m}$, for some value of l. Thus the "Hauptvermutung" would imply the negative answer to the question, proposed by W. Hurewicz,[6] "are two closed manifolds of the same homotopy type necessarily homeomorphic?" In fact two closed simplicial manifolds of the same homotopy type need not even have the same nucleus, which adds interest to the question "is the nucleus of a complex a topological invariant?" This theorem on lens spaces, and a simpler example given in §6, also imply the negative answer to each of the questions Q. 1, $\cdots$, 4, which were left open in S.S., §10.

2. Let $\mathfrak{R}$ be any ring with a unit element e. We recall that an element $a \in \mathfrak{R}$ is said to be *regular* if, and only if, there are elements a' and a'', such that $a'a = aa'' = e$. In this case $a' = a'aa'' = a'' = a^{-1}$, say, and a^{-1} is called the *inverse* of a. The square matrices of a given degree n (i.e. having n rows and columns) with elements in $\mathfrak{R}$, also constitute a ring with a unit element. A regular element in this ring of matrices will be called a *regular matrix*, and we shall denote the inverse of a regular matrix $\mathbf{a}$ by $\mathbf{a}^{-1}$. Let

$$\mathbf{r} = \begin{Vmatrix} \mathbf{a}_1 & \mathbf{b}_1 \\ 0 & \mathbf{b}_2 \end{Vmatrix}$$

be a regular matrix of degree $m + n$, where $\mathbf{a}_1$ is a square matrix of degree m, $\mathbf{b}_2$ is a square matrix of degree n and 0 represents a rectangular array of zeros (i.e. each of its elements is the zero in $\mathfrak{R}$). Further let $\mathbf{b}_2$ have a left inverse, meaning an n-rowed square matrix $\mathbf{c}$, such that $\mathbf{cb}_2 = \mathbf{e}_n$ where $\mathbf{e}_n$ denotes the unit matrix of degree n.

LEMMA 1. *Under these conditions* $\mathbf{a}_1$ *and* $\mathbf{b}_2$ *are regular matrices and*

$$\mathbf{r}^{-1} = \begin{Vmatrix} \mathbf{a}_1^{-1} & \mathbf{a}_2' \\ 0 & \mathbf{b}_2^{-1} \end{Vmatrix}$$

where $\mathbf{a}_2' = -\mathbf{a}_1^{-1}\mathbf{b}_1\mathbf{b}_2^{-1}$.

Let

$$\mathbf{r}^{-1} = \begin{Vmatrix} \mathbf{a}_1' & \mathbf{a}_2' \\ \mathbf{b}_1' & \mathbf{b}_2' \end{Vmatrix}$$

---

[4] See H. Seifert and W. Threlfall, Lehrbuch der Topologie, Leipzig (1933), p. 215.

[5] J. W. Alexander, Proc. Nat. Acad. of Sciences, 10 (1924), 99–101; Seifert and Threlfall, loc. cit., p. 279. This contradicts J. H. C. Whitehead, Quarterly Journal of Math. Oxford series, 10 (1939), 81–3, which will be referred to as Q.J. In Q.J. I overlooked the importance of the additional row, which the n^{th} incidence matrix of an n-dimensional complex receives by reason of an elementary expansion of order $n + 1$. This error was pointed out to me by Shaun Wylie.

[6] Loc. cit., p. 125.

where $\mathbf{a}_1'$ and $\mathbf{b}_2'$ are square matrices of degrees m and n respectively. Then the relation $\mathbf{r}\mathbf{r}^{-1} = \mathbf{e}_{m+n}$ implies $\mathbf{a}_1\mathbf{a}_1' + \mathbf{b}_1\mathbf{b}_1' = \mathbf{e}_m$, $\mathbf{a}_1\mathbf{a}_2' + \mathbf{b}_1\mathbf{b}_2' = 0$, $\mathbf{b}_2\mathbf{b}_1' = 0$ and $\mathbf{b}_2\mathbf{b}_2' = \mathbf{e}_n$. Since $\mathbf{c}\mathbf{b}_2 = \mathbf{b}_2\mathbf{b}_2' = \mathbf{e}_n$, it follows that $\mathbf{b}_2$ is regular, that $\mathbf{b}_2' = \mathbf{b}_2^{-1}$, and also that $\mathbf{b}_1' = \mathbf{b}_2^{-1}\mathbf{b}_2\mathbf{b}_1' = 0$. Therefore $\mathbf{a}_1\mathbf{a}_1' = \mathbf{e}_m$, and the relation $\mathbf{r}^{-1}\mathbf{r} = \mathbf{e}_{m+n}$ implies $\mathbf{a}_1'\mathbf{a}_1 = \mathbf{e}_m$, whence $\mathbf{a}_1$ is regular and $\mathbf{a}_1' = \mathbf{a}_1^{-1}$. Finally $\mathbf{a}_2' = -\mathbf{a}_1^{-1}\mathbf{b}_1\mathbf{b}_2' = -\mathbf{a}_1^{-1}\mathbf{b}_1\mathbf{b}_2^{-1}$ and the proof is complete.

Let $\mathfrak{G}$ be a specified sub-group of the multiplicative group of regular elements in $\mathfrak{R}$. By an *elementary $\mathfrak{G}$-matrix* or simply an *elementary matrix* if no confusion is to be feared, we shall mean a square matrix of any degree n, which can be transformed into the unit matrix $\mathbf{e}_n$ by a finite sequence of elementary transformations, each of which consists either of

1. *multiplying each element in a row, from the left, or a column, from the right, by the same (arbitrary) element of $\mathfrak{G}$, or*
2. *adding a left multiple of any row to some other row, or of*
3. *adding a right multiple of any column to some other column, the multiplier in each of the last two cases being an arbitrary element of $\mathfrak{R}$.*

The inverse of each of these elementary transformations is clearly a transformation of the same type. Therefore the elementary matrices of a given degree constitute a sub-group of the multiplicative group of regular matrices of that degree. As with unimodular matrices of integers it may be verified that, if $\mathbf{a}$ is an elementary matrix of degree n and $\mathbf{r}$ is any matrix with n rows (columns), then $\mathbf{r}$ can be transformed into $\mathbf{ar}$($\mathbf{ra}$) by a sequence of the elementary transformations 1 and 2 (1 and 3). Notice also that, if $\| m_{ij} \|$ is any square matrix of integers, whose determinant is unity, and if $g \in \mathfrak{G}$, then $\| m_{ij}\, g \|$ is an elementary matrix. If $\mathfrak{G}$ contains the regular element $-e$ the same applies to $\| m_{ij}\, g \|$ if $| m_{ij} | = -1$. In any case

$$\begin{Vmatrix} 0 & -\mathbf{e}_n \\ \mathbf{e}_n & 0 \end{Vmatrix}$$

is an elementary $\mathfrak{G}$-matrix of degree $2n$, since, replacing e by 1, its determinant is $(-1)^{n^2+n} = 1$.

LEMMA 2. *If $\mathbf{a}$ is a regular matrix of degree n, then*

$$\mathbf{r} = \begin{Vmatrix} \mathbf{a} & 0 \\ \mathbf{b} & \mathbf{a}^{-1} \end{Vmatrix}$$

is an elementary matrix, where $\mathbf{b}$ is any square matrix of degree n.

By reiterating transformations of the form $\rho_i \to \rho_i + \lambda\rho_j$, where ρ_i and ρ_j $(i \neq j)$ stand for rows and $\lambda \in \mathfrak{R}$, we can transform $\mathbf{r}$ into

$$\mathbf{r}' = \begin{Vmatrix} \mathbf{a} & 0 \\ \mathbf{b} + \mathbf{ca} & \mathbf{a}^{-1} \end{Vmatrix},$$

where $\mathbf{c}$ is any square matrix of degree n. Taking $\mathbf{c} = (\mathbf{e}_n - \mathbf{b})\mathbf{a}^{-1}$ we have

$$\mathbf{r}' = \begin{Vmatrix} \mathbf{a} & 0 \\ \mathbf{e}_n & \mathbf{a}^{-1} \end{Vmatrix},$$

and $\mathbf{r}'$ is transformable into the elementary matrix

$$\begin{Vmatrix} 0 & -\mathbf{e}_n \\ \mathbf{e}_n & 0 \end{Vmatrix}$$

by the sequence of transformations

$$\begin{Vmatrix} \mathbf{a} & 0 \\ \mathbf{e}_n & \mathbf{a}^{-1} \end{Vmatrix} \rightarrow \begin{Vmatrix} 0 & -\mathbf{e}_n \\ \mathbf{e}_n & \mathbf{a}^{-1} \end{Vmatrix} \rightarrow \begin{Vmatrix} 0 & -\mathbf{e}_n \\ \mathbf{e}_n & 0 \end{Vmatrix},$$

in which, first, the bottom rectangle $\| \mathbf{e}_n , \mathbf{a}^{-1} \|$ is multiplied on the left by $-\mathbf{a}$ and added to the top, and then $\| 0, -\mathbf{e}_n \|$ is multiplied by $\mathbf{a}^{-1}$ and added to the bottom. The lemma now follows from the fact that a matrix which is transformable by the elementary transformations into an elementary matrix is itself elementary.

Let us introduce the convention that a matrix may have m rows (columns) and no columns (rows), for any $m = 0, 1, \cdots$. If a matrix $\mathbf{a}$ has m rows and p columns, and if $\mathbf{b}$ has n rows and q columns, where $m, n, p, q \geqq 0$, then, as usual, the product $\mathbf{ab}$, with $m \geqq 0$ rows and $q \geqq 0$ columns, shall exist if, and only if, $p = n$. If $m, q > 0$, $p = n = 0$, then $\mathbf{ab}$ shall be the zero matrix with m rows and q columns. In formulae which involve $\mathbf{e}_k$, with $k \geqq 0$, $\mathbf{e}_0$ is to be the empty matrix with no rows or columns, which we include among the elementary and regular matrices. With these conventions let $\mathbf{r}^n, \cdots, \mathbf{r}^1$ be any set of matrices, such that $\mathbf{r}^p$ has α_p rows and α_{p-1} columns ($p = 1, \cdots, n$; $\alpha_p \geqq 0$). We proceed to define two kinds of equivalence between $\mathbf{r}^n, \cdots, \mathbf{r}^1$ and a set of matrices $\mathbf{s}^n, \cdots, \mathbf{s}^1$, which satisfy a similar condition. That is to say, we can form the product $\mathbf{s}^{p+1}\mathbf{s}^p$ for each $p = 1, \cdots, n-1$, though $\mathbf{s}^p$ need not have the same shape as $\mathbf{r}^p$. We first of all allow ourselves to border $\mathbf{r}^p$ with a new last row and column, which have the common element e in the bottom right hand corner, at the same time adding a final row of zeros to $\mathbf{r}^{p-1}$, if $p > 1$, or an empty row if $\alpha_{p-2} = 0$, and a final column of zeros to $\mathbf{r}^{p+1}$, if $p < n$, or an empty column if $\alpha_{p+1} = 0$ (if $\alpha_p = \alpha_{p-1} = 0$, then $\mathbf{r}^p$ will be transformed into the matrix with the single element e). We shall call this operation an *elementary expansion* of the matrices $\mathbf{r}^n, \cdots, \mathbf{r}^1$. We shall also call it a *primary expansion* of $\mathbf{r}^p$, and shall describe the corresponding expansions of $\mathbf{r}^{p\pm1}$, as *secondary expansions*. In order to avoid the qualifications "if $p > 1$" and "if $p < n$" we shall assume that $\mathbf{r}^{n+1}$ and $\mathbf{r}^0$ exist, where $\mathbf{r}^{n+1}$ has no rows and α_n columns and $\mathbf{r}^0$ has α_0 rows and no columns. A primary expansion shall never be applied to $\mathbf{r}^0$, and shall only be applied to $\mathbf{r}^{n+1}$ when special permission is granted. The inverse of an elementary expansion will be called an *elementary contraction*. Two sets of

matrices $\mathbf{r}^n, \cdots, \mathbf{r}^1$ and $\mathbf{s}^n, \cdots, \mathbf{s}^1$, such that the products $\mathbf{r}^{p+1}\mathbf{r}^p$ and $\mathbf{s}^{p+1}\mathbf{s}^p$ exist for each $p = 1, \cdots, n-1$, will be described as *L-equivalent* if and only if

$$\bar{\mathbf{s}}^p = \mathbf{a}_p \bar{\mathbf{r}}^p \mathbf{a}_{p-1}^{-1} \qquad (p = 1, \cdots, n), \tag{2.1}$$

where $\bar{\mathbf{r}}^n, \cdots, \bar{\mathbf{r}}^1$ and $\bar{\mathbf{s}}^n, \cdots, \bar{\mathbf{s}}^1$ are obtained from $\mathbf{r}^n, \cdots, \mathbf{r}^1$ and $\mathbf{s}^n, \cdots, \mathbf{s}^1$ by sequences of elementary expansions, and $\mathbf{a}_p$ is an elementary matrix of the appropriate degree, for each $p = 0, \cdots, n$. Two sets of matrices, $\mathbf{r}^n, \cdots, \mathbf{r}^1$ and $\mathbf{s}^n, \cdots, \mathbf{s}^1$, will be described as L^*-equivalent if, and only if, they satisfy the same condition, except that $\mathbf{a}_p$ may be any regular matrix of the appropriate degree. These relations are obviously symmetric between the two sets of matrices.

We now assume that the group $\mathfrak{G}$ contains $-e$, in which case two rows or columns of a given matrix may be interchanged by a sequence of elementary transformations. This being so, if we apply a sequence of transformations to the matrices $\mathbf{r}^n, \cdots, \mathbf{r}^1$, which consist of elementary expansions and contractions, and transformations of the form (2.1), a simple induction shows that all the expansions may be applied first,[7] then a transformation of the form (2.1) and finally the contractions. Hence it follows that the relations of L-equivalence and L^*-equivalence are equivalence relations in the technical sense, meaning that they are symmetric and transitive.

LEMMA 3. *If $\mathbf{r}^n, \cdots, \mathbf{r}^1$ and $\mathbf{s}^n, \cdots, \mathbf{s}^1$ are L^*-equivalent, then they can be expanded into sets of matrices $\bar{\mathbf{r}}^n, \cdots, \bar{\mathbf{r}}^1$ and $\bar{\mathbf{s}}^n, \cdots, \bar{\mathbf{s}}^1$, which are related by equations of the form* (2.1), *subject to the condition that $\mathbf{a}_0, \cdots, a_{n-1}$ shall be elementary matrices. Thus the sets $\mathbf{a}\bar{\mathbf{r}}^n, \bar{\mathbf{r}}^{n-1}, \cdots, \bar{\mathbf{r}}^1$ and $\bar{\mathbf{s}}^n, \cdots, \bar{\mathbf{s}}^1$ are L-equivalent, where $\mathbf{a}$ is some regular matrix.*

After initial expansions we assume that $\mathbf{r}^n, \cdots, \mathbf{r}^1$ and $\mathbf{s}^n, \cdots, \mathbf{s}^1$ are themselves related by transformations of the form (2.1), in which $\mathbf{a}_0, \cdots, \mathbf{a}_{p-1}$ are elementary matrices ($0 \leqq p \leqq n$), this last condition being vacuous if $p = 0$. If $p = n$ there is nothing to prove. Otherwise we expand $\mathbf{r}^{p+2}$, $\mathbf{r}^{p+1}$, $\mathbf{r}^p$ into

$$\bar{\mathbf{r}}^{p+2} = \| \mathbf{r}^{p+2}, 0 \|, \qquad \bar{\mathbf{r}}^{p+1} = \begin{Vmatrix} \mathbf{r}^{p+1} & 0 \\ 0 & \mathbf{e}_{\alpha_p} \end{Vmatrix}, \qquad \bar{\mathbf{r}}^p = \begin{Vmatrix} \mathbf{r}^p \\ 0 \end{Vmatrix},$$

and write

$$\bar{\mathbf{s}}^{p+2} = \mathbf{a}_{p+2} \| \mathbf{r}^{p+2}, 0 \| \begin{Vmatrix} \mathbf{a}_{p+1}^{-1} & 0 \\ 0 & \mathbf{a}_p \end{Vmatrix} = \| \mathbf{s}^{p+2}, 0 \|$$

$$\bar{\mathbf{s}}^{p+1} = \begin{Vmatrix} \mathbf{a}_{p+1} & 0 \\ 0 & \mathbf{a}_p^{-1} \end{Vmatrix} \begin{Vmatrix} \mathbf{r}^{p+1} & 0 \\ 0 & \mathbf{e}_{\alpha_p} \end{Vmatrix} \begin{Vmatrix} \mathbf{a}_p^{-1} & 0 \\ 0 & \mathbf{a}_p \end{Vmatrix} = \begin{Vmatrix} \mathbf{s}^{p+1} & 0 \\ 0 & \mathbf{e}_{\alpha_p} \end{Vmatrix},$$

$$\bar{\mathbf{s}}^p = \begin{Vmatrix} \mathbf{a}_p & 0 \\ 0 & \mathbf{a}_p^{-1} \end{Vmatrix} \begin{Vmatrix} \mathbf{r}^p \\ 0 \end{Vmatrix} \mathbf{a}_{p-1}^{-1} = \begin{Vmatrix} \mathbf{s}^p \\ 0 \end{Vmatrix}.$$

The lemma now follows from lemma 2 and induction on p.

[7] Cf. the preliminary argument in theorem 1 below.

We now define what we shall temporarily call *EL*-equivalence (extended *L*-equivalence). The conditions for *EL*-equivalence are the same as for *L*-equivalence except that we add empty matrices $\mathbf{r}^{n+1}, \mathbf{r}^{n+2}, \cdots$ to a given set $\mathbf{r}^n, \cdots, \mathbf{r}^1$, and allow a primary expansion of $\mathbf{r}^p$, even if $p > n$, where initially $\mathbf{r}^p$ has no rows or columns if $p > n + 1$, and has no rows and α_n columns if $p = n + 1$. If $p > n$ a primary expansion of $\mathbf{r}^p$ will be called an *extended expansion* of the set $\mathbf{r}^n, \cdots, \mathbf{r}^1$. It is clear that *EL*-equivalence, like *L*-equivalence and for a similar reason, is an equivalence relation in the technical sense.

THEOREM 1. *If the matrices* $\mathbf{r}^n, \cdots, \mathbf{r}^1$ *and* $\mathbf{s}^n, \cdots, \mathbf{s}^1$ *are EL-equivalent, then they are L-equivalent.*

To say that $\mathbf{r}^n, \cdots, \mathbf{r}^1$ and $\mathbf{s}^n, \cdots, \mathbf{s}^1$ are *EL*-equivalent is to say that matrices $\bar{\mathbf{r}}^m, \cdots, \bar{\mathbf{r}}^1$ and $\bar{\mathbf{s}}^m, \cdots, \bar{\mathbf{s}}^1$ are related by a transformation of the form (2.1), with n replaced by $m \geqq n$, where $\bar{\mathbf{r}}^m, \cdots, \bar{\mathbf{r}}^1$ and $\bar{\mathbf{s}}^m, \cdots, \bar{\mathbf{s}}^1$ are obtained from $\mathbf{r}^n, \cdots, \mathbf{r}^1$ and $\mathbf{s}^n, \cdots, \mathbf{s}^1$ by expansions and extended expansions and $\mathbf{a}_0, \cdots, \mathbf{a}_m$ are elementary matrices. The theorem will follow from induction on m when we have proved that $\mathbf{r}^n, \cdots, \mathbf{r}^1$ and $\mathbf{s}^n, \cdots, \mathbf{s}^1$ are *EL*-equivalent under a transformation in which no primary expansion is applied to $\mathbf{r}^p$ if $p > m - 1$, unless $m = n$, in which case there is nothing to prove. I say that, in the given sequence of expansions, the primary expansions of $\mathbf{r}^m$, if there are any, may be applied last of all. For if $p < m - 1$ a primary expansion of $\mathbf{r}^p$ is obviously interchangeable with a primary expansion of $\mathbf{r}^m$. The effect of expanding $\mathbf{r}^{m-1}$ before, instead of after $\mathbf{r}^m$, is to interchange the rows which are added by the given secondary and primary expansions of $\mathbf{r}^{m-1}$, and also the corresponding columns of $\mathbf{r}^m$. Therefore we may postpone the primary expansions of $\mathbf{r}^m$ till the final stages of the expansion and, if necessary, include certain permutations of the rows of $\bar{\mathbf{r}}^{m-1}$ and the columns of $\bar{\mathbf{r}}^m$ in the transformation (2.1). So we have to prove that matrices $\mathbf{r}'^{m-1}, \cdots, \bar{\mathbf{r}}^1$ and $\mathbf{s}'^{m-1}, \cdots, \bar{\mathbf{s}}^1$ are *L*-equivalent if $\bar{\mathbf{r}}^m, \bar{\mathbf{r}}^{m-1}, \bar{\mathbf{r}}^{m-2}, \cdots, \bar{\mathbf{r}}^1$ and $\bar{\mathbf{s}}^m, \bar{\mathbf{s}}^{m-1}, \bar{\mathbf{s}}^{m-2}, \cdots, \bar{\mathbf{s}}^1$ are related by a transformation of the form (2.1), in which $\mathbf{a}_0, \cdots, \mathbf{a}_m$ are elementary matrices, where

$$\begin{aligned} \bar{\mathbf{r}}^m &= \|\, 0, \mathbf{e}_k \,\|, \qquad \bar{\mathbf{r}}^{m-1} = \left\| \begin{matrix} \mathbf{r}'^{m-1} \\ 0 \end{matrix} \right\|, \\ \bar{\mathbf{s}}^m &= \|\, 0, \mathbf{e}_l \,\|, \qquad \bar{\mathbf{s}}^{m-1} = \left\| \begin{matrix} \mathbf{s}'^{m-1} \\ 0 \end{matrix} \right\|, \end{aligned} \tag{2.2}$$

and $k = l$, since the matrices $\bar{\mathbf{r}}^m$ and $\bar{\mathbf{s}}^m$ have the same shape. This being so, we have

$$\begin{aligned} \mathbf{a}\,\|\, 0, \mathbf{e}_k \,\| \left\| \begin{matrix} \mathbf{p}_1 & \mathbf{q}_1 \\ \mathbf{p}_2 & \mathbf{q}_2 \end{matrix} \right\| &= \|\, 0, \mathbf{a} \,\| \left\| \begin{matrix} \mathbf{p}_1 & \mathbf{q}_1 \\ \mathbf{p}_2 & \mathbf{q}_2 \end{matrix} \right\| \\ &= \|\, \mathbf{a}\mathbf{p}_2, \mathbf{a}\mathbf{q}_2 \,\| \\ &= \|\, 0, \mathbf{e}_k \,\|, \end{aligned}$$

where $\mathbf{a} = \mathbf{a}_m$ and

$$\mathbf{a}_{m-1}^{-1} = \left\| \begin{matrix} \mathbf{p}_1 & \mathbf{q}_1 \\ \mathbf{p}_2 & \mathbf{q}_2 \end{matrix} \right\|,$$

$\mathbf{q}_2$ being a square matrix of degree k.

Therefore $\mathbf{p}_2 = 0$ and $\mathbf{q}_2 = \mathbf{a}^{-1}$. Since $\mathbf{a}^{-1}$ is regular it follows from lemma 1 that $\mathbf{a}_{m-1}$ is of the form

$$\mathbf{a}_{m-1} = \left\| \begin{matrix} \mathbf{b} & \mathbf{c} \\ 0 & \mathbf{a} \end{matrix} \right\|,$$

whence

$$\mathbf{a}_{m-1}\bar{\mathbf{r}}_{m-1} = \left\| \begin{matrix} \mathbf{b}\mathbf{r}'^{m-1} \\ 0 \end{matrix} \right\|.$$

It follows from an argument in the proof of lemma 2 that

$$\left\| \begin{matrix} \mathbf{b} & \mathbf{c} - \mathbf{c}\mathbf{a}^{-1}\mathbf{a} \\ 0 & \mathbf{a} \end{matrix} \right\| = \left\| \begin{matrix} \mathbf{b} & 0 \\ 0 & \mathbf{a} \end{matrix} \right\|$$

is an elementary matrix. Therefore, writing $n = m - 1$ and omitting the primes and dashes, the theorem will follow when we have proved that two sets of matrices $\mathbf{r}^n, \cdots, \mathbf{r}^1$ and $\mathbf{b}\mathbf{r}^n, \cdots, \mathbf{r}^1$ are L-equivalent if $\mathbf{b}$ is a square matrix, such that

$$\left\| \begin{matrix} \mathbf{b} & 0 \\ 0 & \mathbf{a} \end{matrix} \right\|$$

is an elementary matrix, where $\mathbf{a}$ is some elementary matrix.[8] To prove this we first expand $\mathbf{r}^n$, transforming $\mathbf{r}^n$ and $\mathbf{r}^{n-1}$ into

$$\left\| \begin{matrix} \mathbf{r}^n & 0 \\ 0 & \mathbf{e}_k \end{matrix} \right\| \quad \text{and} \quad \left\| \begin{matrix} \mathbf{r}^{n-1} \\ 0 \end{matrix} \right\|,$$

where k is the degree of $\mathbf{a}$. Since $\mathbf{a}$ is an elementary matrix, so obviously is

$$\left\| \begin{matrix} \mathbf{e}_\alpha & 0 \\ 0 & \mathbf{a} \end{matrix} \right\| \qquad (\alpha = \alpha_n).$$

[8] See note II_1 in §6 below.

Therefore $\mathbf{r}^n, \cdots, \mathbf{r}^1$ are L-equivalent to $\bar{\mathbf{r}}^n, \bar{\mathbf{r}}^{n-1}, \cdots, \mathbf{r}^1$, where

$$\begin{aligned}
\bar{\mathbf{r}}^n &= \left\| \begin{matrix} \mathbf{b} & 0 \\ 0 & \mathbf{a} \end{matrix} \right\| \left\| \begin{matrix} \mathbf{r}^n & 0 \\ 0 & \mathbf{e}_k \end{matrix} \right\| \left\| \begin{matrix} \mathbf{e}_\alpha & 0 \\ 0 & \mathbf{a}^{-1} \end{matrix} \right\| \\
&= \left\| \begin{matrix} \mathbf{br}^n & 0 \\ 0 & \mathbf{a} \end{matrix} \right\| \left\| \begin{matrix} \mathbf{e}_\alpha & 0 \\ 0 & \mathbf{a}^{-1} \end{matrix} \right\| \\
&= \left\| \begin{matrix} \mathbf{br}^n & 0 \\ 0 & \mathbf{e}_k \end{matrix} \right\|, \\
\bar{\mathbf{r}}^{n-1} &= \left\| \begin{matrix} \mathbf{e}_\alpha & 0 \\ 0 & \mathbf{a} \end{matrix} \right\| \left\| \begin{matrix} \mathbf{r}^{n-1} \\ 0 \end{matrix} \right\| \\
&= \left\| \begin{matrix} \mathbf{r}^{n-1} \\ 0 \end{matrix} \right\|.
\end{aligned}$$

But $\bar{\mathbf{r}}^n, \bar{\mathbf{r}}^{n-1}$ contract into $\mathbf{br}^n, \mathbf{r}^{n-1}$, and the proof is complete.

Two sets of matrices, $\mathbf{r}^n, \cdots, \mathbf{r}^1$ and $\mathbf{s}^n, \cdots, \mathbf{s}^1$, will be described as *$L^m$-equivalent*, for any $m \geqq 0$, if, and only if, they satisfy the conditions for EL-equivalence, except that, in addition to expansions and extended expansions, we may add a final row of zeros to $\mathbf{r}^k$ and a final column of zeros to $\mathbf{r}^{k+1}$, for any $k > m$, on the understanding that, if $\alpha_{k-1} = 0$, then $\left\| \begin{matrix} \mathbf{r}^k \\ 0 \end{matrix} \right\|$ is an empty matrix having $\alpha_k + 1$ rows and no columns, with a similar convention concerning $\| \mathbf{r}^{k+1}, 0 \|$ if $\alpha_{k+1} = 0$. Thus, in passing from $\mathbf{r}^n, \cdots, \mathbf{r}^1$ to $\mathbf{s}^n, \cdots, \mathbf{s}^1$ by a series of elementary transformations, one may, at any stage, apply such a transformation or its inverse. Notice that L-equivalence implies L^m-equivalence for every value of m, and that L^m-equivalence implies L^q-equivalence if $m > q$. Clearly L^m-equivalence is also a symmetric and transitive relation. We shall prove that, if $m \geqq n$, then L^m-equivalence is the same as L^*-equivalence, subject to the following condition on the ring $\mathfrak{R}$. If

$$\mathbf{a} \left\| \begin{matrix} 0 & \mathbf{e}_k \\ 0 & 0 \end{matrix} \right\| = \left\| \begin{matrix} 0 & \mathbf{e}_l \\ 0 & 0 \end{matrix} \right\| \mathbf{b}, \tag{2.3}$$

where $\mathbf{a}$ and $\mathbf{b}$ are regular matrices of appropriate degrees, then $k = l$. This condition is satisfied[9] if there is a homomorphism, $\phi(\mathfrak{R}) = \mathfrak{R}_0$, of $\mathfrak{R}$ on a commutative ring $\mathfrak{R}_0$, which contains at least two elements. For the determinant of any square matrix, whose elements are in $\mathfrak{R}_0$, can be calculated in the ordinary way, and hence the rank of any matrix. If $\mathbf{s} = \mathbf{arb}$, $\mathbf{r} = \mathbf{a'sb'}$, where $\mathbf{a}$, $\mathbf{b}$, $\mathbf{a'}$, $\mathbf{b'}$, $\mathbf{r}$ and $\mathbf{s}$ are matrices with elements in $\mathfrak{R}_0$, it follows from the standard argument that $\mathbf{r}$ and $\mathbf{s}$ have the same rank. Since $\mathfrak{R}_0$ contains at least two ele-

[9] See note II_2 in §6 below.

ments and since $\phi(e)\phi(x) = \phi(ex) = \phi(x)$ for any $\phi(x) \in \mathfrak{R}_0$, it follows that $\mathfrak{R}_0$ has a unit element, namely $\phi(e)$. Therefore $\phi(e)$ is not nilpotent, whence a determinant with $\phi(e)$ down the main diagonal, and zeros everywhere else, is not zero. On replacing e, a_{ij} and $b_{\lambda\mu}$, in (2.3), by $\phi(e)$, $\phi(a_{ij})$ and $\phi(b_{\lambda\mu})$, where $\mathbf{a} = \| a_{ij} \|$, $\mathbf{b} = \| b_{\lambda\mu} \|$, it follows from equality of rank that $k = l$.

THEOREM 2. *Subject to the above condition on $\mathfrak{R}$, if two sets of matrices $\mathbf{r}^n, \cdots, \mathbf{r}^1$ and $\mathbf{s}^n, \cdots, \mathbf{s}^1$ are L^n-equivalent, then they are L^*-equivalent. Conversely, L^*-equivalence implies L^m-equivalence for every value of m.*

Let $\mathbf{r}^n, \cdots, \mathbf{r}^1$ and $\mathbf{s}^n, \cdots, \mathbf{s}^1$ be L^n-equivalent. Then $\mathbf{r}^n, \cdots, \mathbf{r}^1$ and $\mathbf{s}^n, \cdots, \mathbf{s}^1$ can be expanded, by elementary and extended expansions and transformations of the form $\mathbf{r}^{k+1}$, $\mathbf{r}^k \to \| \mathbf{r}^{k+1}, 0 \|$, $\left\| \begin{matrix} \mathbf{r}^k \\ 0 \end{matrix} \right\|$ $(k > n)$, into sets $\bar{\mathbf{r}}^q, \cdots, \bar{\mathbf{r}}^1$ and $\bar{\mathbf{s}}^q, \cdots, \bar{\mathbf{s}}^1$ $(q \geqq n)$, such that $\bar{\mathbf{s}}^p = \mathbf{a}_p \bar{\mathbf{r}}^p \mathbf{a}_{p-1}^{-1}$ $(p = 1, \cdots, q)$, where $\mathbf{a}_0, \cdots, \mathbf{a}_q$ are elementary matrices. As in the proof of theorem 1 we may suppose that the primary expansions of $\mathbf{r}^n, \cdots, \mathbf{r}^1$ and $\mathbf{s}^n, \cdots, \mathbf{s}^1$ are applied first in the expansions $\mathbf{r}^p \to \bar{\mathbf{r}}^p$ and $\mathbf{s}^p \to \bar{\mathbf{s}}^p$. Without altering the notation, let us assume that they have already been applied. Then, after rearranging the rows and columns of $\bar{\mathbf{r}}^{n+1}$ and $\bar{\mathbf{s}}^{n+1}$ if necessary, we have

$$\bar{\mathbf{r}}^{n+1} = \left\| \begin{matrix} 0 & \mathbf{e}_k \\ 0 & 0 \end{matrix} \right\|, \quad \bar{\mathbf{r}}^n = \left\| \begin{matrix} \mathbf{r}^n \\ 0 \end{matrix} \right\|$$

$$\bar{\mathbf{s}}^{n+1} = \left\| \begin{matrix} 0 & \mathbf{e}_l \\ 0 & 0 \end{matrix} \right\|, \quad \bar{\mathbf{s}}^n = \left\| \begin{matrix} \mathbf{s}^n \\ 0 \end{matrix} \right\| \qquad (k, l \geqq 0),$$

where the bottom rectangles of zeros in $\bar{\mathbf{r}}^n$ and $\bar{\mathbf{s}}^n$ contain k and l rows respectively, and the bottom rectangles of zeros in $\bar{\mathbf{r}}^{n+1}$ and $\bar{\mathbf{s}}^{n+1}$ may be empty. If $k = l = 0$ the matrices $\mathbf{r}^n, \cdots, \mathbf{r}^1$ and $\mathbf{s}^n, \cdots, \mathbf{s}^1$ are L-equivalent. Otherwise it follows from the relation $\mathbf{a}_{n+1}\bar{\mathbf{r}}^{n+1} = \bar{\mathbf{s}}^{n+1}\mathbf{a}_n$ and the condition (2.3) that $k = l$, and we have

$$\left\| \begin{matrix} \mathbf{b}_1 & \mathbf{c}_1 \\ \mathbf{b}_2 & \mathbf{c}_2 \end{matrix} \right\| \left\| \begin{matrix} 0 & \mathbf{e}_k \\ 0 & 0 \end{matrix} \right\| = \left\| \begin{matrix} 0 & \mathbf{e}_k \\ 0 & 0 \end{matrix} \right\| \left\| \begin{matrix} \mathbf{p}_1 & \mathbf{q}_1 \\ \mathbf{p}_2 & \mathbf{q}_2 \end{matrix} \right\|, \tag{2.4}$$

where

$$\left\| \begin{matrix} \mathbf{b}_1 & \mathbf{c}_1 \\ \mathbf{b}_2 & \mathbf{c}_2 \end{matrix} \right\| = \mathbf{a}_{n+1}, \quad \left\| \begin{matrix} \mathbf{p}_1 & \mathbf{q}_1 \\ \mathbf{p}_2 & \mathbf{q}_2 \end{matrix} \right\| = \mathbf{a}_n,$$

$\mathbf{b}_1$ and $\mathbf{q}_2$ being square matrices of degree k. Also

$$\left\| \begin{matrix} \mathbf{s}^n \\ 0 \end{matrix} \right\| = \left\| \begin{matrix} \mathbf{p}_1 & \mathbf{q}_1 \\ \mathbf{p}_2 & \mathbf{q}_2 \end{matrix} \right\| \left\| \begin{matrix} \mathbf{r}^n \\ 0 \end{matrix} \right\| \mathbf{a}_{n-1}^{-1} = \left\| \begin{matrix} \mathbf{p}_1 \mathbf{r}^n \mathbf{a}_{n-1}^{-1} \\ \mathbf{p}_2 \mathbf{r}^n \mathbf{a}_{n-1}^{-1} \end{matrix} \right\|.$$

Since $k = l$ it follows that $\mathbf{s}^n = \mathbf{p}_1\mathbf{r}^n\mathbf{a}_{n-1}^{-1}$, and we have only to show that the matrix $\mathbf{p}_1$ is regular. From (2.4) we have

$$\left\|\begin{matrix}0 & \mathbf{b}_1\\ 0 & \mathbf{b}_2\end{matrix}\right\| = \left\|\begin{matrix}\mathbf{p}_2 & \mathbf{q}_2\\ 0 & 0\end{matrix}\right\|,$$

whence $\mathbf{b}_2 = 0$, $\mathbf{p}_2 = 0$, $\mathbf{b}_1 = \mathbf{q}_2$. Let

$$\left\|\begin{matrix}\mathbf{b}_1 & \mathbf{c}_1\\ \mathbf{b}_2 & \mathbf{c}_2\end{matrix}\right\|^{-1} = \left\|\begin{matrix}\mathbf{b}_1' & \mathbf{b}_2'\\ \mathbf{c}_1' & \mathbf{c}_2'\end{matrix}\right\|.$$

Since $\mathbf{b}_2 = 0$, we have

$$\left\|\begin{matrix}\mathbf{b}_1' & \mathbf{b}_2'\\ \mathbf{c}_1' & \mathbf{c}_2'\end{matrix}\right\|\left\|\begin{matrix}\mathbf{b}_1 & \mathbf{c}_1\\ \mathbf{b}_2 & \mathbf{c}_2\end{matrix}\right\| = \left\|\begin{matrix}\mathbf{b}_1'\mathbf{b}_1 & *\\ * & *\end{matrix}\right\| = \mathbf{e}_r$$

for some $r > 0$. Therefore $\mathbf{b}_1$ has a left inverse. Since $\mathbf{p}_2 = 0$, $\mathbf{q}_2 = \mathbf{b}_1$, and since

$$\left\|\begin{matrix}\mathbf{p}_1 & \mathbf{q}_1\\ 0 & \mathbf{q}_2\end{matrix}\right\|$$

is a regular matrix, it follows from lemma 1 that $\mathbf{p}_1$ is a regular matrix, and the first part of the theorem is established.

Conversely, let $\mathbf{r}^n, \cdots, \mathbf{r}^1$ and $\mathbf{s}^n, \cdots, \mathbf{s}^1$ be L^*-equivalent and, given $m \geqq 0$, let us introduce empty matrices $\mathbf{r}^{m+1}, \cdots, \mathbf{r}^{n+1}$ and $\mathbf{s}^{m+1}, \cdots, \mathbf{s}^{n+1}$ if $m \geqq n$. Then the sets $\mathbf{r}^k, \cdots, \mathbf{r}^1$ and $\mathbf{s}^k, \cdots, \mathbf{s}^1$ are obviously L^*-equivalent, where $k = \max(m, n) + 1$, and by lemma 3 there are sets $\bar{\mathbf{r}}^k, \cdots, \bar{\mathbf{r}}^1$ and $\bar{\mathbf{s}}^k, \cdots, \bar{\mathbf{s}}^1$, which are L-equivalent to $\mathbf{r}^k, \cdots, \mathbf{r}^1$ and $\mathbf{s}^k, \cdots, \mathbf{s}^1$ respectively, such that $\bar{\mathbf{s}}^k = \mathbf{a}\bar{\mathbf{r}}^k$, $\bar{\mathbf{s}}^p = \bar{\mathbf{r}}^p$ for $p = 1, \cdots, k-1$, where $\mathbf{a}$ is a regular matrix. Let

$$\bar{\mathbf{r}}^k = \left\|\begin{matrix}\bar{\mathbf{r}}^k\\ 0\end{matrix}\right\|, \qquad \bar{\mathbf{s}}^k = \left\|\begin{matrix}\bar{\mathbf{s}}^k\\ 0\end{matrix}\right\|,$$

where the number of zero rows indicated by 0 is the same as the number of rows in $\bar{\mathbf{r}}^k$ and $\bar{\mathbf{s}}^k$. Then

$$\bar{\mathbf{s}}^k = \left\|\begin{matrix}\mathbf{a} & 0\\ 0 & \mathbf{a}^{-1}\end{matrix}\right\|\bar{\mathbf{r}}^k,$$

and the theorem follows from lemma 2.

COROLLARY. *If* $\mathbf{r}^n, \cdots, \mathbf{r}^1$ *and* $\mathbf{s}^n, \cdots, \mathbf{s}^1$ *are* L^n*-equivalent, then they are* L^m*-equivalent for every* $m \geqq 0$.

In the second part of theorem 2 it follows from the argument in theorem 1, which is referred to in the first part of theorem 2, that the transformations $\mathbf{r}^k$, $\mathbf{s}^k \to \left\|\begin{matrix}\mathbf{r}^k\\ 0\end{matrix}\right\|, \left\|\begin{matrix}\mathbf{s}^k\\ 0\end{matrix}\right\|$ may be applied before any primary expansions of $\mathbf{r}^k, \cdots, \mathbf{r}^1$ an $\mathbf{s}^k, \cdots, \mathbf{s}^1$. Therefore, taking $m \geqq n$, we have the addendum to theorem 2:

ADDENDUM. *If* $\mathbf{r}^n, \cdots, \mathbf{r}^1$ *and* $\mathbf{s}^n, \cdots, \mathbf{s}^1$ *are* L^**-equivalent, then* $\mathbf{r}^{m+1}, \cdots, \mathbf{r}^1$ *and* $\mathbf{s}^{m+1}, \cdots, \mathbf{s}^1$ *are* L*-equivalent, for any* $m \geqq n$, *where* $\mathbf{r}^{m+1}$ *and* $\mathbf{s}^{m+1}$ *are certain empty or zero matrices if* $m = n$, *and* $\mathbf{r}^p$ *and* $\mathbf{s}^p$ $(p = n+1, \cdots, m+1)$ *are certain empty matrices if* $m > n$ $(\alpha_{m+1} \geqq 0, \alpha_m = \cdots = \alpha_{n+1} = 0)$.

We now assume that $\mathbf{r}^p\mathbf{r}^{p-1} = 0$ for each $p = 2, \cdots, n$, where 0 stands for an empty matrix if either $\alpha_p = 0$ or $\alpha_{p-2} = 0$. This condition is obviously invariant under all our elementary transformations and under transformations of the form (2.1). On this assumption, we shall prove a lemma, which brings the elementary transformations by which L-equivalence is defined into closer relation with elementary deformations of complexes. Let us extend the terms *elementary expansion* and *primary expansion* so as to include any transformation of the form

$$
\begin{aligned}
&\mathbf{r}^{p+1} \to \| \mathbf{r}^{p+1}, \quad 0 \| \\
&\mathbf{r}^p \to \begin{Vmatrix} \mathbf{r}^p & 0 \\ \mathbf{a} & g \end{Vmatrix} = \begin{Vmatrix} \mathbf{r}^p & 0 \\ 0 & e \end{Vmatrix} \begin{Vmatrix} \mathbf{e}_\alpha & 0 \\ \mathbf{a} & g \end{Vmatrix} \qquad (\alpha = \alpha_{p-1}) \\
&\mathbf{r}^{p-1} \to \begin{Vmatrix} \mathbf{r}^{p-1} \\ \mathbf{b} \end{Vmatrix} = \begin{Vmatrix} \mathbf{e}_\alpha & 0 \\ -g^{-1}\mathbf{a} & g^{-1} \end{Vmatrix} \begin{Vmatrix} \mathbf{r}^{p-1} \\ 0 \end{Vmatrix},
\end{aligned} \tag{2.5}
$$

where $\mathbf{a}$ stands for any row $\| a_1, \cdots, a_\alpha \|$ $(a_\lambda \in \mathfrak{R})$, $g \in \mathfrak{G}$ and $\mathbf{b}$ stands for the row $-g^{-1}\mathbf{a}\mathbf{r}^{p-1}$, of which the j^{th} element is

$$
-g^{-1} \sum_{\lambda=1}^{\alpha} a_\lambda r_{\lambda j}^{p-1} \qquad (j = 1, \cdots, \alpha_{p-2}; \mathbf{r}^{p-1} = \| \mathbf{r}_{\lambda j}^{p-1} \|).
$$

The inverse of such a transformation will be called an *elementary contraction*. In this definition it is to be understood that the new rows and columns are indicated as the last rows and columns in (2.5) merely for convenience of notation. Thus the removal of the μ^{th} column and the k^{th} row of $\mathbf{r}^p$ in case $\mathbf{r}_{k\mu}^p \in G$, $r_{j\mu}^p = 0$ if $j \neq k$, together with the μ^{th} row of $\mathbf{r}^{p-1}$ and the k^{th} column of $\mathbf{r}^{p+1}$ will be called an elementary contraction. Notice that, if $r_{k\mu}^p = g$, $r_{j\mu}^p = 0$ $(j \neq k)$, where $g \in \mathfrak{G}$, or if g is any regular element, then the μ^{th} row of $\mathbf{r}^{p-1}$ has the form indicated above in consequence of $\mathbf{r}^p\mathbf{r}^{p-1} = 0$ and the k^{th} column of $\mathbf{r}^{p+1}$ consists entirely of zeros since $\mathbf{r}^{p+1}\mathbf{r}^p = 0$.

Let $\mathbf{r}^n, \cdots, \mathbf{r}^1$ and $\mathbf{s}^n, \cdots, \mathbf{s}^1$ be L-equivalent sets of matrices $(\mathbf{r}^p\mathbf{r}^{p-1} = 0, \mathbf{s}^p\mathbf{s}^{p-1} = 0)$. We add empty matrices $\mathbf{r}^{n+1}$ and $\mathbf{s}^{n+1}$ to these sets, on the understanding that a primary expansion of $\mathbf{r}^{n+1}$ is to be regarded as an elementary expansion, not as an extended expansion.

LEMMA 4. *The matrices* $\mathbf{r}^{n+1}, \cdots, \mathbf{r}^1$ *can be transformed into* $\mathbf{s}^{n+1}, \cdots, \mathbf{s}^1$ *by a sequence of elementary expansions and contractions.*

First let $\mathbf{r}^p = \mathbf{s}^p$ for $p = 1, \cdots, n-1$ and let $\mathbf{s}^n = \mathbf{a}\mathbf{r}^n$, where $\mathbf{a}$ is an elementary matrix. Then $\mathbf{r}^n$ can be transformed into $\mathbf{s}^n$ by a sequence of elementary transformations of the form $\rho_i \to g\rho_i + \lambda\rho_j$ $(i \neq j)$, where $g \in G$, $\lambda \in \mathfrak{R}$ and ρ_k denotes the k^{th} row of a given matrix. Taking $i = 1$, $j = 2$, for sim-

plicity of notation, this transformation is the resultant of the elementary expansion and contraction

$$\mathbf{r}^{n+1} \to \| e, \quad -g, \quad -\lambda, \quad 0 \| \to \mathbf{r}^{n+1},$$

$$\mathbf{r}^n \to \begin{Vmatrix} g\rho_1 + \lambda\rho_2 \\ \rho_1 \\ \vdots \\ \rho_{\alpha_n} \end{Vmatrix} \to \begin{Vmatrix} g\rho_1 + \lambda\rho_2 \\ \rho_2 \\ \vdots \\ \rho_{\alpha_n} \end{Vmatrix},$$

in which the row $g\rho_1 + \lambda\rho_2$ is created and ρ_1 is then deleted. Thus the special case of the lemma is established.

In general, we assume, as in lemma 3, that, after certain initial expansions, the two sets of matrices are already related by equations of the form (2.1), in which $\mathbf{a}_0, \cdots, \mathbf{a}_n$ are elementary matrices. Let $\mathbf{a}_0, \cdots, \mathbf{a}_{p-1}$ be unit matrices, for some value of p $(0 \leqq p < n)$, this condition being vacuous if $p = 0$. Then we expand $\mathbf{r}^{p+2}$, $\mathbf{r}^{p+1}$, $\mathbf{r}^p$ and $\mathbf{s}^{p+2}$, $\mathbf{s}^{p+1}$, $\mathbf{s}^p$ into the matrices $\bar{\mathbf{r}}^q$, $\bar{\mathbf{s}}^q$ $(q = p, p+1, p+2)$, where

$$\bar{\mathbf{r}}^{p+2} = \| 0, \quad \mathbf{r}^{p+2} \|, \qquad \bar{\mathbf{r}}^{p+1} = \begin{Vmatrix} \mathbf{e}_{\alpha_p} & -\mathbf{a}_p \\ 0 & \mathbf{r}^{p+1} \end{Vmatrix}, \qquad \bar{\mathbf{r}}^p = \begin{Vmatrix} \mathbf{a}_p\mathbf{r}^p \\ \mathbf{r}^p \end{Vmatrix} = \begin{Vmatrix} \mathbf{s}^p \\ \mathbf{r}^p \end{Vmatrix},$$

$$\bar{\mathbf{s}}^{p+2} = \| 0, \quad \mathbf{s}^{p+2} \|, \qquad \bar{\mathbf{s}}^{p+1} = \begin{Vmatrix} -\mathbf{a}_p^{-1} & \mathbf{e}_{\alpha_p} \\ \mathbf{s}^{p+1} & 0 \end{Vmatrix}, \qquad \bar{\mathbf{s}}^p = \begin{Vmatrix} \mathbf{s}^p \\ \mathbf{a}_p^{-1}\mathbf{s}^p \end{Vmatrix} = \begin{Vmatrix} \mathbf{s}^p \\ \mathbf{r}^p \end{Vmatrix} = \bar{\mathbf{r}}^p.$$

If $\mathbf{a}$ and $\mathbf{b}$ are elementary matrices, so obviously is $\begin{Vmatrix} \mathbf{a} & 0 \\ 0 & \mathbf{b} \end{Vmatrix}$. Therefore it follows from an argument in the proof of lemma 2 that

$$\bar{\mathbf{a}}_{p+1} = \begin{Vmatrix} -\mathbf{a}_p^{-1} & 0 \\ \mathbf{s}^{p+1} & \mathbf{a}_{p+1} \end{Vmatrix}$$

is an elementary matrix, and since $\mathbf{s}^{p+1}\mathbf{a}_p = \mathbf{a}_{p+1}\mathbf{r}^{p+1}$ we have

$$\bar{\mathbf{a}}_{p+1}\bar{\mathbf{r}}^{p+1} = \begin{Vmatrix} -\mathbf{a}_p^{-1} & \mathbf{e}_{\alpha_p} \\ \mathbf{s}^{p+1} & 0 \end{Vmatrix} = \bar{\mathbf{s}}^{p+1}.$$

Also it may be verified that

$$\bar{\mathbf{a}}_{p+1}^{-1} = \begin{Vmatrix} -\mathbf{a}_p & 0 \\ \mathbf{r}^{p+1} & \mathbf{a}_{p+1}^{-1} \end{Vmatrix},$$

and since $\mathbf{r}^{p+2}\mathbf{r}^{p+1} = 0$ it follows that

$$\begin{aligned} \mathbf{a}_{p+2}\bar{\mathbf{r}}^{p+2}\bar{\mathbf{a}}_{p+1}^{-1} &= \| 0, \quad \mathbf{a}_{p+2}\mathbf{r}^{p+2} \| \begin{Vmatrix} -\mathbf{a}_p & 0 \\ \mathbf{r}^{p+1} & \mathbf{a}_{p+1}^{-1} \end{Vmatrix} \\ &= \| 0, \quad \mathbf{a}_{p+2}\mathbf{r}^{p+2}\mathbf{a}_{p+1}^{-1} \| \\ &= \bar{\mathbf{s}}^{p+2}. \end{aligned}$$

The conditions are now as before, with p replaced by $p + 1$. Therefore the lemma follows from the special case and induction on p.

We conclude this section by recovering, in a modified form, a result due to Reidemeister,[10] concerning three given matrices $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$, which satisfy certain special conditions. First, the ring $\mathfrak{R}$, to which the elements of $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ belong shall be commutative. Secondly, $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ shall be of the form $\mathbf{a} = \| a_\lambda^i \|$, $\mathbf{b} = \| b_\rho^\lambda \|$ and $\mathbf{c} = \| c_t^\rho \|$ $(i = 1, \cdots, n;\ \lambda = 1, \cdots, m;\ \rho = 1, \cdots, p;\ t = 1, \cdots, q)$, where the superscripts indicate the rows, and finally we assume that $m - n = p - q = r \geqq 0$. Then, using a notation which is familiar in tensor analysis,[11] we define

$$A_{\lambda_1 \cdots \lambda_n} = e_{i_1 \cdots i_n} a_{\lambda_1}^{i_1} \cdots a_{\lambda_n}^{i_n},$$

$$B_{\lambda_1 \cdots \lambda_n}^{\rho_1 \cdots \rho_q} = \frac{1}{r!} e_{\lambda_1 \cdots \lambda_n \mu_1 \cdots \mu_r} e^{\rho_1 \cdots \rho_q \sigma_1 \cdots \sigma_r} b_{\sigma_1}^{\mu_1} \cdots b_{\sigma_r}^{\mu_r},$$

$$C^{\rho_1 \cdots \rho_q} = e^{t_1 \cdots t_q} c_{t_1}^{\rho_1} \cdots c_{t_q}^{\rho_q},$$

with the convention that $B = e$, the unit element in $\mathfrak{R}$, if $r = 0$. Stated in words, $A_{\lambda_1 \cdots \lambda_n}$ and $C^{\rho_1 \cdots \rho_q}$ are the determinants $| a_{\lambda_j}^i |$ and $| c_t^{\rho_s} |$ $(i, j = 1, \cdots, n;\ s, t = 1, \cdots, q)$ while $\pm B_{\lambda_1 \cdots \lambda_n}^{\rho_1 \cdots \rho_q}$ is the determinant of the square matrix obtained from $\| b_\rho^\lambda \|$ by striking out the rows $b^{\lambda_1}, \cdots, b^{\lambda_n}$ and the columns $b_{\rho_1}, \cdots, b_{\rho_q}$, unless $\lambda_i = \lambda_j$ or $\rho_s = \rho_t$ for some $i \neq j$ or $s \neq t$, in which case $B_{\lambda_1 \cdots \lambda_n}^{\rho_1 \cdots \rho_q} = 0$. We assume that

$$A_{\lambda_1 \cdots \lambda_n} C^{\rho_1 \cdots \rho_q} = \kappa B_{\lambda_1 \cdots \lambda_n}^{\rho_1 \cdots \rho_q}, \tag{2.6}$$

for some $\kappa \in \mathfrak{R}$, and all values of the indices λ_i, ρ_s. Then it follows from the tensorial character of the components A, B and C that this condition is invariant under transformations of the form

$$\mathbf{a}' = \mathbf{fag}^{-1}, \qquad \mathbf{b}' = \mathbf{gbh}^{-1}, \qquad \mathbf{c}' = \mathbf{hck}^{-1},$$

where $\mathbf{f}$, $\mathbf{g}$, $\mathbf{h}$ and $\mathbf{k}$ are regular matrices of appropriate degrees. Moreover, the condition will be satisfied after such a transformation if κ is replaced by $\kappa' = \pi\kappa$, where π is a product of the determinants $f = | f_j^i |$, $g = | g_\mu^\lambda |$, $h = | h_\sigma^\rho |$, $k = | k_t^s |$ and their inverses (actually $\pi = f^{-1} g h^{-1} k$). Clearly the determinant of a regular matrix is a regular element of $\mathfrak{R}$, and the determinant of an elementary $\mathfrak{G}$-matrix is an element of $\mathfrak{G}$. Therefore the multiplier π is a regular element, and $\pi \in \mathfrak{G}$ if $\mathbf{f}$, $\mathbf{g}$, $\mathbf{h}$ and $\mathbf{k}$ are elementary matrices. Finally it is easy to see that the condition (2.6) is invariant under an elementary expansion (not an extended expansion) of the matrices $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$, the factor κ being unaltered. Therefore we have the lemma:

LEMMA 5. *If the matrices* $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ *satisfy a condition of the form* (2.6), *so do any matrices* $\mathbf{a}'$, $\mathbf{b}'$, $\mathbf{c}'$, *which are* L^**-equivalent to* $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, *and with* κ *replaced by* $\kappa' = \pi\kappa$,

[10] Journal für die Math. r. u. a. (loc. cit.), §6. See also W. Franz, idem., 176 (1937), 113-34.

[11] See, for example, O. Veblen, Invariants of quadratic differential forms, Cambridge (1933), chap. 1.

where π is a regular element of $\mathfrak{A}$. If $\mathbf{a}'$, $\mathbf{b}'$, $\mathbf{c}'$ are L-equivalent to $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, then we may take $\pi \in \mathfrak{G}$.

Clearly κ is uniquely determined by (2.6), unless there is a non-zero $\lambda \in \mathfrak{R}$, such that $\lambda B_{\lambda_1}^{\rho_1 \cdots \rho_q}{}_{\lambda_n} = 0$ for all values of the indices. The condition that there shall, or shall not, exist such a λ is obviously an invariant of L^*-equivalence. Therefore, if κ is uniquely determined, so is κ', and π is uniquely determined unless κ is a 0-divisor, including 0 itself among the 0-divisors. In particular, κ is uniquely determined if $r = 0$.

3. In this and the next section we study the relation between the algebra in §2 and the topology of finite, connected complexes. We first correct the error in Q.J. Let K_1^n and K_2^p ($n \geqq p$) be two finite, connected simplicial complexes and let $\mathbf{r}^n, \cdots, \mathbf{r}^1$ and $\mathbf{s}^p, \cdots, \mathbf{s}^1$ be incidence matrices for K_1^n and K_2^p, which are determined by Reidemeister's process. The elements of $\mathbf{r}^j$ and $\mathbf{s}^k$ are in the group rings, $\mathfrak{R}_1$, and $\mathfrak{R}_2$, of K_1^n and K_2^p. The (integral) group ring $\mathfrak{R}$, of any group $\mathfrak{H}$, satisfies the condition concerning k and l in (2.3). For a homomorphism of $\mathfrak{R}$ on the coefficient ring of $\mathfrak{R}$, in this case the ring of integers, is determined by writing $x = 1$ for each $x \in \mathfrak{H}$. Returning to K_1^n and K_2^p, if $n > p$ we add empty matrices $\mathbf{s}^n, \cdots, \mathbf{s}^{p+1}$ to the set $\mathbf{s}^p, \cdots, \mathbf{s}^1$ and agree that, if $p < q \leqq n$, then a primary expansion of $\mathbf{s}^q$ shall be an elementary expansion, not an extended expansion. Let K_1^n and K_2^p be of the same homotopy type. Then $\pi_1(K_1^n)$ is isomorphic to $\pi_1(K_2^p)$, and by a *special isomorphism* of $\mathfrak{R}_1$ on $\mathfrak{R}_2$ we shall mean one which is determined by an isomorphism of $\pi_1(K_1^n)$ on $\pi_1(K_2^p)$. In the definition of elementary $\mathfrak{G}$-matrices, and L-equivalence of the incidence matrices for K_1^n, we now take $\mathfrak{G}$ to be the group consisting of all the elements $\pm x$, where $x \in \pi_1(K_1^n)$.

Let $\mathfrak{R}_1$ and $\mathfrak{R}_2$ be two isomorphic rings and let $\psi(\mathfrak{R}_1) = \mathfrak{R}_2$ be an isomorphism of $\mathfrak{R}_1$ on $\mathfrak{R}_2$. If $\mathbf{r} = \| r_{i\lambda} \|$, where $r_{i\lambda} \in \mathfrak{R}_1$, and if $r'_{i\lambda} = \psi(r_{i\lambda})$, we shall denote the matrix $\| r'_{i\lambda} \|$ by $\psi(\mathbf{r})$. We shall describe two sets of matrices $\mathbf{r}^n, \cdots, \mathbf{r}^1$ and $\mathbf{s}^n, \cdots, \mathbf{s}^1$, whose elements belong to $\mathfrak{R}_1$ and $\mathfrak{R}_2$ respectively, as *X-equivalent under ψ* ($X = L$ or L^*) if, and only if, the matrices $\psi(\mathbf{r}^n), \cdots, \psi(\mathbf{r}^1)$ and $\mathbf{s}^n, \cdots, \mathbf{s}^1$ are X-equivalent. The terms *EL*-equivalent and L^m-equivalent under ψ will have a similar meaning. What does follow from Q.J., is that the incidence matrices of K_1^n and K_2^p are *EL*-equivalent under some special isomorphism $\psi(\mathfrak{R}_1) = \mathfrak{R}_2$ if K_1^n and K_2^p have the same nucleus. It then follows from S.S., theorems 17 and 13 (pp. 277, 269) that the incidence matrices of K_1^n and K_2^p are L^n-equivalent if K_1^n and K_2^p are of the same homotopy type. Hence, theorems 1 and 2 lead to the following theorem, which is also a corollary of theorems 6 and 8 in §4 below.

THEOREM 3. *If K_1^n and K_2^p have the same nucleus their incidence matrices are L-equivalent under some special isomorphism $\psi(\mathfrak{R}_1) = \mathfrak{R}_2$. If K_1^n and K_2^p are of the same homotopy type their incidence matrices are L^*-equivalent under a special isomorphism $\psi(\mathfrak{R}_1) = \mathfrak{R}_2$.*

In either case it follows from induction on the number of elementary trans-

formations by which one passes from K_1^n to K_2^p that the isomorphism ψ may be taken from the class which is determined by some homotopy isomorphism[12] $f(K_1^n) \subset K_2^p$.

Theorem 3 states that the geometrical operations described in S.S. can be copied algebraically by the operations described in §2 above. We now turn to the converse question: "to what extent can the algebraic transformations be copied geometrically?" In order to study this question we introduce what we call membrane complexes, which are, so to speak, more elastic than simplicial complexes. Before defining a membrane complex, we modify the definition of a simple membrane, given in S.S., p. 263, by including among the simple membranes, first an n-element E^n, and secondly any stellar subdivision $\sigma\mathbf{E}^n$, as a simple membrane, where $\mathbf{E}^n$ is any simple membrane. If $\mathbf{E}^n = E^n$, an n-element, then its boundary[13] $F(\mathbf{E}^n)$, shall be the identical map of $\dot{E}^n$ on itself. Let $F(\mathbf{E}^n)$ be a simplicial map $f(S^{n-1}) \subset K$, where $\mathbf{E}^n$ is a given membrane, and let (A, a) be an elementary sub-division of $\mathbf{E}^n$. If the simplex A is internal to $\mathbf{E}^n$ then $F(\sigma\mathbf{E}^n) = f(S^{n-1})$, where σ is the subdivision (A, a). If $A \subset F(\mathbf{E}^n)$, let σ_1 be the sub-division of S^{n-1} in which all the simplexes in $f^{-1}(A)$ are starred. Then $F(\sigma\mathbf{E}^n)$ shall be the simplicial map $f'(\sigma_1 S^{n-1})$, in which $f'(b) = a$ if b is one of the new vertices introduced by σ_1, and $f'(b) = f(b)$ if b is one of the original vertices in S^{n-1}. By reiterating this construction we define $F(\sigma\mathbf{E}^n)$, where σ is any stellar sub-division. It is obvious that this extension of the definition does not invalidate any of the results in S.S.

An n-dimensional membrane complex $\mathbf{K}^n$, will be defined by induction on n. A 0-dimensional membrane is a 0-simplex and a 0-dimensional membrane complex is a collection of 0-simplexes. We assume that an $(n-1)$-dimensional complex $\mathbf{K}^{n-1}$, has been defined ($n \geqq 1$), and that it is a simplicial complex of at most $n-1$ dimensions, whose simplexes are oriented and grouped together to form certain oriented simple membranes, which we shall describe as the *cells* of $\mathbf{K}^{n-1}$. Then an *n-dimensional membrane complex* is to be a complex of the form[14]

$$\mathbf{K}^n = \mathbf{K}^{n-1} + \mathbf{E}_1^n + \cdots + \mathbf{E}_{a_n}^n,$$

where $\mathbf{E}_i^n$ is an oriented, simple membrane, whose boundary is a simplicial, spherical map $F(\mathbf{E}_i^n) \subset \mathbf{K}^{n-1}$. Thus $\mathbf{K}^n$ is a simplicial complex K^n, whose simplexes are grouped together in a particular way to form the cells of $\mathbf{K}^n$.

---

[12] See note III$_1$.

[13] As in S.S., $F(\mathbf{E}^p)$ will always denote the "spherical" boundary of a simple membrane $\mathbf{E}^p$. $F(\mathbf{E}^p)$ is a (simplicial) map, not a complex. We shall use $\dot{E}^n$ to denote the boundary of an n-element E^n. The cycle in $\mathbf{K}^{p-1}$, defined below, which is determined by $F(\mathbf{E}_i^p)$, will be denoted by $\partial\mathbf{E}_i^p$. In fact $\partial\mathbf{E}_i^p$ will only be used for membranes in a universal covering complex of a given complex.

[14] As a matter of convention we allow this set of membranes to be empty, in which case $\mathbf{K}^n = \mathbf{K}^{n-1}$. Thus $\mathbf{K}^n$ stands for a complex of at most n dimensions in the ordinary sense of the word, and it is to be assumed throughout this paper that $n > 2$. As in S.S., it is to be assumed that, in the above expression for $\mathbf{K}^n$, or in any similar expression, none of the membranes has in inner point in common with any of the others or with $\mathbf{K}^{n-1}$.

We shall call K^n, or any stellar sub-division of K^n, a *simplicial sub-division* of $\mathbf{K}^n$ and shall say that K^n covers $\mathbf{K}^n$. We shall use $\mathbf{K}^p$ $(0 \leqq p \leqq n)$ to stand for the membrane complex consisting of the cells in $\mathbf{K}^n$ whose dimensionalities do not exceed p, and K^p will stand for the sub-complex of K^n which covers $\mathbf{K}^p$ (not for the p-dimensional skeleton of K^n). Notice that $F(\mathbf{E}_i^p)$, where $\mathbf{E}_i^p \subset \mathbf{K}^p$, though it is a simplicial map in K^{p-1}, need not cover an exact sum of cells in $\mathbf{K}^{p-1}$.

It follows from the definition in S.S. that a simple membrane is a combinatorial cell,[15] and from an argument in S.S., lemma 5, that, if K^{q+1} is a simplicial subdivision of $\mathbf{K}^{q+1}$, then any p-chain $(p \leqq q)$ in K^{q+1} can be deformed into $\mathbf{K}^q$, with the part in $\mathbf{K}^q$ held fixed. Therefore the p-cells $\mathbf{E}_i^p \subset \mathbf{K}^p$ constitute an homology basis[16] for the p-cycles in any simplicial subdivision of $\mathbf{K}^n$. It is to be understood that any 'given complex' is finite and connected, but we note that these definitions apply equally well to infinite complexes. In particular a universal covering complex $\tilde{\mathbf{K}}^n$, of $\mathbf{K}^n$, is obviously a membrane complex, whose cells are the oriented membranes which cover the membranes of $\mathbf{K}^n$. It is always to be understood that the (oriented) cells of $\tilde{\mathbf{K}}^n$ cover the cells of $\mathbf{K}^n$ positively.

We shall restrict the formal deformations of a membrane complex to deformations in which the membranes may be regarded as undivided cells,[17] and the following arguments show that the main results in S.S. are applicable to the resulting scheme. Let $\mathbf{K}$ be a given membrane complex and let $\mathbf{E}^p$ be a new membrane bounded by a simplicial spherical map $F(\mathbf{E}^p) \subset \mathbf{K}$, which is homotopic to a point in $\mathbf{K}$. Let $f(S^p) \subset \mathbf{K} + \mathbf{E}^p$ be a simplicial map which is *simple* in $\mathbf{E}^p$, meaning that there is a p-simplex $A^p \subset E^p$, such that $f^{-1}(A^p)$ is a single simplex. Let $\mathbf{E}^{p+1}$ be a new membrane which is bounded by $f(S^p)$, and let $\mathbf{K}_1 = \mathbf{K} + \mathbf{E}^{p+1}$. Then the transformation $\mathbf{K} \to \mathbf{K}_1$ will be called an *elementary expansion* of the membrane complex $\mathbf{K}$, and its inverse will be called an *elementary contraction*. By a *formal deformation* D we shall mean a finite sequence of elementary expansions and contractions, whose resultant is a transformation $\mathbf{K} \to D(\mathbf{K})$. Two membrane complexes $\mathbf{K}_0$ and $\mathbf{K}$ will be said to have the same *nucleus* if, and only if, $\mathbf{K} = D(\mathbf{K}_0)$, where D is some formal deformation. Similarly we take over the definition of an m-group from S.S., defining an elementary filling of order p as a transformation of the form $\mathbf{K} \to \mathbf{K} + \mathbf{E}^p$, where $\mathbf{E}^p$ is a new membrane whose boundary is an arbitrary simplicial map in $\mathbf{K}^{p-1}$.

Theorem 5 in S.S. is obviously valid for membrane complexes, on the understanding that $\sigma\mathbf{K}$ is the membrane complex which is obtained from $\mathbf{K}$ by a subdivision σK, where K is a simplicial sub-division of $\mathbf{K}$. Theorems 11 and 14 in S.S. (pp. 264 and 272) are valid for membrane complexes. Indeed the proof of theorem 11 can be considerably simplified. For if $F(\mathbf{E}_i^n) \subset \mathbf{K}$ $(i = 0, 1)$ and

[15] See, for example, P. Alexandroff and H. Hopf, Topologie, Berlin (1935), p. 245.

[16] That is to say, any p-cycle in $\mathbf{K}^n$ is homologous to a cycle which consists of (integral) multiples of these cells, and if such a p-cycle bounds in $\mathbf{K}^n$ it bounds a chain composed of the $(p+1)$-dimensional membranes (cf. Alexandroff and Hopf, loc. cit., pp. 245 et seq.).

[17] Cf. note III_2.

if $F(\mathbf{E}_0^n)$ is homotopic in $\mathbf{K}$ to $F(\mathbf{E}_1^n)$, then there is obviously a spherical map $f(S^n) \subset \mathbf{K} + \mathbf{E}_0^n + \mathbf{E}_1^n$, which is simple both in $\mathbf{E}_0^n$ and in $\mathbf{E}_1^n$. Therefore, assuming first that $\mathbf{E}_i^n$ does not meet $\mathbf{K} + \mathbf{E}_j^n (i, j = 0, 1; j \neq i)$ except in $F(\mathbf{E}_i^n)$, it follows that $\mathbf{K} + \mathbf{E}_i^n$ expands into $\mathbf{K} + \mathbf{E}^{n+1}$, where $F(\mathbf{E}^{n+1}) = f(S^n)$. Therefore

$$\mathbf{K} + \mathbf{E}_1^n = D(\mathbf{K} + \mathbf{E}_0^n) \qquad \text{(rel. } \mathbf{K}\text{)},$$

where 'rel. $\mathbf{K}$' means the same as in S.S. If $\mathbf{E}_0^n$ and $\mathbf{E}_1^n$ have internal intersections, we still have

$$\mathbf{K} + \mathbf{E}_i^n = D(\mathbf{K} + \mathbf{E}_2^n) \qquad \text{(rel. } \mathbf{K}\text{)},$$

where $F(\mathbf{E}_2^n) = F(\mathbf{E}_0^n)$ and $\mathbf{E}_2^n$ does not meet $\mathbf{K} + \mathbf{E}_0^n + \mathbf{E}_1^n$ except in $F(\mathbf{E}_2^n)$. This establishes the generalization of theorem 11, and the generalization of theorem 14 follows from the arguments given in S.S. We shall refer to these generalizations simply as S.S., theorems 11 and 14.

Let $\mathbf{K} \to \mathbf{K}_1 = \mathbf{K} + \mathbf{E}^{n+1}$ be an elementary expansion of $\mathbf{K}$, where the map $F(\mathbf{E}^{n+1}) \subset \mathbf{K} + \mathbf{E}^n$ is simple in the membrane $\mathbf{E}^n$, which does not meet $\mathbf{K}$ except in $F(\mathbf{E}^n) \subset \mathbf{K}$. Then I say that K expands geometrically into K_1, in the sense of S.S., §5 (p. 258), where K_1 is a simplicial sub-division of $\mathbf{K}_1$ and K is the subcomplex of K_1 which covers $\mathbf{K}$. For, first let $\mathbf{E}^n = E^n$, an n-element, and let $A^n \subset E^n$ be an n-simplex, which has but a single original in the map $F(\mathbf{E}^{n+1})$. After a suitable sub-division of E^n, we may assume that all the vertices of A^n are internal to E^n, and the assertion follows from S.S., theorem 1 (p. 250), the corollary to theorem 8 (p. 260) and lemma 6 (p. 265). If $\mathbf{E}^n$ is of the form $E^n + C_f(\dot{E}^n)$, as defined in S.S., p. 263, then $\mathbf{E}^n - \mathbf{A}^n$ contracts into $F(\mathbf{E}^n)$, where $\mathbf{A}^n$ is any open simplex in $\mathbf{E}^n$. For $\mathbf{E}^n = E_0^n + N$, where E_0^n is an n-element which contains[18] E^n in its interior, and N is the set of closed simplexes in $\mathbf{E}^n$, which meet $F(\mathbf{E}^n)$. If $\mathbf{A}^n \epsilon E^n$ it follows that $\mathbf{E}^n - \mathbf{A}^n$ contracts geometrically into N, and, from an argument similar to the proof of theorem 8 in S.S., that N contracts into $F(\mathbf{E}^n)$. If $\mathbf{A}^n \epsilon C_f(\dot{E}^n)$ it follows from an argument similar to one on p. 281 of S.S. that $\mathbf{E}^n - \mathbf{A}^n$ contracts into $F(\mathbf{E}^n)$. Finally, if $\mathbf{E}^n = \sigma\{E^n + C_f(\dot{E}^n)\}$ it follows from what we have just proved and S.S., lemma 7 (p. 267), that $\mathbf{E}^n - \mathbf{A}^n$ contracts geometrically into $F(\mathbf{E}^n)$, where $\mathbf{A}^n$ is any open simplex in $\mathbf{E}^n$. Therefore it follows from S.S., lemma 6, that $\mathbf{K}$ expands geometrically into $\mathbf{K}$, and the assertion is justified in each case. Hence, and from S.S., lemma 5 (p. 265), it follows that, if $\mathbf{K}_0$ and $\mathbf{K}_1$ have the same nucleus or m-group, for a given value of m, then so do any simplicial sub-divisions of $\mathbf{K}_0$ and $\mathbf{K}_1$. The converse, and also the generalization of theorem 17 in S.S. p. 277, follows from the theorem:

THEOREM 4. *If K is a simplicial sub-division of $\mathbf{K}$, then $\mathbf{K} = D(K)$.*

If $\mathbf{K} = D_0(K^*)$, where K^* is any simplicial complex (i.e. a membrane complex whose cells are simplexes), we have $K = D_1(K^*)$, where K is a simplicial subdivision of $\mathbf{K}$, whence $\mathbf{K} = D_0 D_1^{-1}(K)$. Therefore the theorem will follow when

[18] E_0^n may be regarded as a triangulation of $E^n + (\dot{E}^n \times \langle 0, \frac{1}{2}\rangle)$ (Cf. S.S., p. 259).

we have proved that $\mathbf{K} = D_0(K^*)$, where K^* is a simplicial complex. This is trivial if $\mathbf{K}$ consists of a single vertex, and we shall prove it by induction on the number of cells in $\mathbf{K}$. Let $\mathbf{K} = \mathbf{K}_0 + \mathbf{E}_1^n$, where $\mathbf{E}_1^n$ is a cell of maximum dimensionality, and $\mathbf{K}_0$ consists of the remaining cells in $\mathbf{K}$. By the hypothesis of the induction, there is a simplicial complex K_0, such that $K_0 = D_0(\mathbf{K}_0)$, and by S.S., theorem 14, D_0 may be extended to a deformation $\mathbf{K} \to D_1(\mathbf{K}) = K_0 + \mathbf{E}^n = \mathbf{K}_1$, say, where $\mathbf{E}^n = D_1(\mathbf{E}_1^n)$, and $\mathbf{K}_1$ consists of the simplexes in K_0, together with the cell $\mathbf{E}^n$. Let $\mathbf{E}_0^n$ be a new membrane with the same boundary as $\mathbf{E}^n$, which does not meet $\mathbf{K}_1$ except in $F(\mathbf{E}_0^n) \subset K_0^{n-1}$, and let $\mathbf{A}^n \epsilon \mathbf{E}_0^n$ be an open n-simplex, whose closure is internal to $\mathbf{E}_0^n$. Let $K^* = K_0 + \mathbf{E}_0^n$ be the simplicial complex, whose cells are the simplexes in K_0 and in $\mathbf{E}_0^n$, and let $\mathbf{E}^{n+1}$ be a new membrane, whose boundary is the n-sphere[19] $\mathbf{E}^n - \mathbf{E}_0^n$. Then $F(\mathbf{E}^{n+1})$ is simple in $\mathbf{A}^n$, whence $\mathbf{K}_1$ expands first into $\mathbf{K}_1 + \mathbf{E}_0^n - \mathbf{A}^n$, and then into $\mathbf{K}_1 + \mathbf{E}^{n+1}$. But $F(\mathbf{E}^{n+1})$ is also simple in $\mathbf{E}^n$, whence $\mathbf{K}_1 + \mathbf{E}^{n+1}$ contracts into K^*, and the theorem is established.

We now proceed to the algebraic description of a (membrane) complex $\mathbf{K}^n$. Let $\tilde{\mathbf{K}}^n$ be a universal covering complex of $\mathbf{K}^n$, and $u(\tilde{\mathbf{K}}^n) = \mathbf{K}^n$ a locally $(1 - 1)$ map of $\tilde{\mathbf{K}}^n$ on $\mathbf{K}^n$. Let $\mathbf{E}^0 = \mathbf{E}_1^0$, say, be an arbitrary vertex of $\mathbf{K}^n$ and let $\tilde{\mathbf{E}}^0 \epsilon \tilde{\mathbf{K}}^n$ be an arbitrary vertex in $u^{-1}(\mathbf{E}^0)$. We take $\mathbf{E}^0$ as the base point for $\pi_1(\mathbf{K}^n)$, and identify each element in the covering group[20] of $\tilde{\mathbf{K}}^n$ with the element of $\pi_1(\mathbf{K}^n)$ to which it corresponds in the isomorphism determined by $\mathbf{E}^0$ and $\tilde{\mathbf{E}}^0$. Then an arbitrary vertex in $u^{-1}(\mathbf{E}^0)$ will be denoted by $x\tilde{\mathbf{E}}^0$, where $x \epsilon \pi_1(\mathbf{K}^n)$. Let $\mathbf{V}$ be a (connected) tree which contains $\mathbf{K}^0$ and is a sub-complex of $\mathbf{K}^1$, and let $x\tilde{\mathbf{V}}$ be the component of $u^{-1}(\mathbf{V}) \subset \tilde{\mathbf{K}}^n$ which contains $x\tilde{\mathbf{E}}^0$. Let $\mathbf{E}_1^1, \cdots, \mathbf{E}_k^1$ be the oriented 1-cells of $\mathbf{K}^1$, if there are any, which are not in $\mathbf{V}$, and let $x\tilde{\mathbf{E}}_i^1$ be the 1-cell of $\tilde{\mathbf{K}}^n$ which covers $\mathbf{E}_i^1$ positively and whose initial point is on $x\tilde{\mathbf{V}}$. In calculating the incidence matrices $\mathbf{r}^2$, $\mathbf{r}^1$ for $\mathbf{K}^n$, we shall treat $\mathbf{V}$ as if it were a single point, ignoring the 1-cells in $\mathbf{V}$ and the vertices other than $\mathbf{E}^0$. On this understanding, we shall take[21] $\tilde{\mathbf{V}} = 1\tilde{\mathbf{V}}$ and $\tilde{\mathbf{E}}_i^1 = 1\tilde{\mathbf{E}}_i^1$ as basis elements for the 0-chains and 1-chains in $\tilde{\mathbf{K}}^n$. Similarly we shall treat $\mathbf{V}$ as if it were a single point in setting up a system of generators and relations for $\pi_1(\mathbf{K}^n)$. Thus any segment $s \subset \mathbf{K}^1$, which begins and ends on $\mathbf{V}$, is to be automatically transformed into the circuit $s_0 + s - s_1$, where s_0 and s_1 are segments on $\mathbf{V}$ which join $\mathbf{E}^0$ to the first and last points of s, and $s_i (i = 0, 1)$ is non-singular unless it degenerates into $\mathbf{E}^0$. Such a circuit is determined, up to homotopic deformation over $\mathbf{K}^1$, by an expression of the form

$$s = \epsilon_1 \mathbf{E}_{i_1}^1 + \cdots + \epsilon_q \mathbf{E}_{i_q}^1 \qquad (\epsilon_\lambda = \pm 1), \tag{3.1}$$

in which the summation is non-commutative and the last vertex of $\epsilon_\lambda \mathbf{E}_{i_\lambda}^1$ is joined to the first vertex of $\epsilon_{\lambda+1} \mathbf{E}_{i_{\lambda+1}}^1$ by a segment on $\mathbf{V}$. The corresponding segment

[19] Cf. S.S., p. 264.

[20] I.e. the group of homeomorphisms (covering transformations) $t(\tilde{\mathbf{K}}^n) = \tilde{\mathbf{K}}^n$, such that $ut = u$.

[21] We now denote the unit element in any group, and in its group ring, by 1.

in $\tilde{\mathbf{K}}^n$, which starts at $\tilde{\mathbf{V}}$, is given by an expression of the form

$$\tilde{\mathfrak{s}} = \epsilon_1 x_1 \tilde{\mathbf{E}}^1_{i_1} + \cdots + \epsilon_q x_q \tilde{\mathbf{E}}^1_{i_q}, \tag{3.2}$$

where $x_\lambda \in \pi_1(\mathbf{K}^n)$. The circuit (3.1) and the segment (3.2) are also represented by the product

$$x = a_{i_1}^{\epsilon_1} \cdots a_{i_q}^{\epsilon_q}, \tag{3.3}$$

where a_i is the (free) generator of the free group $\mathfrak{F}_k = \pi_1(\mathbf{K}^1)$, which corresponds to the 'circuit' $\mathbf{E}^1_i$. Let $\mathfrak{R}$ be the group ring of $\mathbf{K}^n$, and let

$$C^1 = r_1 \tilde{\mathbf{E}}^1_1 + \cdots + r_k \tilde{\mathbf{E}}^1_k \qquad (r_i \in \mathfrak{R}) \tag{3.4}$$

be the chain which is obtained by simplifying (3.2), with commutative addition. Since $\tilde{\mathfrak{s}}$ is an unbroken segment we have

$$\partial C^1 = (x - 1)\tilde{\mathbf{V}}, \tag{3.5}$$

where ∂C denotes the boundary of a given chain C. In particular $\partial \tilde{\mathbf{E}}^1_i = (a_i - 1)\tilde{\mathbf{V}}$, whence $\mathbf{r}^1$ is the matrix with a single column and $a_i \neq 1$ in the i^{th} row, if $k > 0$. If $k = 0$, then $\mathbf{r}^1$ has one column and no rows.

We now calculate the coefficients x_λ in (3.2) in terms of the product (3.3). Let $\tilde{\mathfrak{s}}_\lambda$ be the part of $\tilde{\mathfrak{s}}$ which is given by $y_\lambda = a_{i_1}^{\epsilon_1} \cdots a_{i_\lambda}^{\epsilon_\lambda}$, with $\tilde{\mathfrak{s}}_0 = \tilde{\mathbf{E}}^0$, $y_0 = 1$. Then the last vertex of $\tilde{\mathfrak{s}}_{\lambda-1}$, and hence the first vertex of $\epsilon_\lambda x_\lambda \tilde{\mathbf{E}}^1_i$, are on $y_{\lambda-1}\tilde{\mathbf{V}}$. Therefore, if $\epsilon_\lambda = +1$, the first vertex of $x_\lambda \tilde{\mathbf{E}}^1_{i_\lambda}$ is on $y_{\lambda-1}\tilde{\mathbf{V}}$, whence $x_\lambda = y_{\lambda-1}$. If $\epsilon_\lambda = -1$ the first vertex of $x_\lambda \tilde{\mathbf{E}}^1_{i_\lambda}$ is on $y_{\lambda-1}a_{i_\lambda}^{-1}\tilde{\mathbf{V}}$, in which case $x_\lambda = y_{\lambda-1}a_{i_\lambda}^{-1}$. Therefore[22]

$$\begin{aligned} x_\lambda &= a_{i_1}^{\epsilon_1} \cdots a_{i_{\lambda-1}}^{\epsilon_{\lambda-1}} \qquad \text{if } \epsilon_\lambda = +1 \\ &= a_{i_1}^{\epsilon_1} \cdots a_{i_\lambda}^{\epsilon_\lambda} \qquad \text{if } \epsilon_\lambda = -1. \end{aligned} \tag{3.6}$$

Notice that, in either case,

$$\begin{aligned} \epsilon_\lambda x_\lambda \partial \tilde{\mathbf{E}}^1_{i_\lambda} &= y_{\lambda-1}(a_{i_\lambda}^{\epsilon_\lambda} - 1)\tilde{\mathbf{V}} \\ &= (y_\lambda - y_{\lambda-1})\tilde{\mathbf{V}}. \end{aligned}$$

Therefore, if ∂ is an abstract operator such that $\partial \tilde{\mathbf{E}}^1_i = (a_i - 1)\tilde{\mathbf{V}}$ and $\partial(r_1 C^1_1 + r_2 C^1_2) = r_1 \partial C^1_1 + r_2 \partial C^1_2$, for any pair of chains C^1_1, C^1_2, and elements $r_1, r_2 \subset \mathfrak{R}$, then (3.5) follows by a purely formal calculation from (3.2) and

[22] In formal calculations which involve (3.6) one must remember the ambiguity in the notation, which is referred to in note III$_5$. Thus (3.3), as an alternative expression for (3.1) or (3.2), is simply a product of the generators. On the other hand the coefficients in (3.2) are elements of $\pi_1(\mathbf{K}^n)$, and x_λ may be replaced by any product of the generators which represents the same element. For example, it may be verified formally that $C^1 = 0$ if $x = 1$ in $\mathfrak{F}_k$, where x and C^1 are given by (3.3) and (3.4), and hence that, if $x_2 = yx_1y^{-1}$, where $x_1 = 1$ in $\pi_1(\mathbf{K}^n)$, and if C^1_i is the chain which corresponds to x_i $(i = 1, 2)$, then $C^1_2 = yC^1_1$ (cf. (4.2), below).

(3.3). Notice also that

$$r_1(a_1 - 1) + \cdots + r_k(a_k - 1) = x - 1 \tag{3.7}$$

in consequence of (3.4) and (3.5).

Let $\mathbf{E}_1^2, \cdots, \mathbf{E}_{\alpha_2}^2$ be the 2-cells of $\mathbf{K}^n$, if there are any, and let $f_\lambda(S_\lambda^1) \subset \mathbf{K}^1$ be a map which is equivalent to the spherical boundary $F(\mathbf{E}_\lambda^2)$. If $f_\lambda(S_\lambda^1) \subset \mathbf{K}^1 - \mathbf{K}^0$ for some value of λ, then $f_\lambda(S_\lambda^1)$ is internal to some 1-cell in $\mathbf{K}^1$, since it is connected, and is therefore homotopic, in $\mathbf{K}^1$, to a point. If not, we choose a base point $p_\lambda \in S_\lambda^1$, such that $f_\lambda(p_\lambda) = \mathbf{E}_\lambda^0 \in \mathbf{V}$, where $\mathbf{E}_\lambda^0$ is any vertex of $\mathbf{K}^0$. In either case $F(\mathbf{E}_\lambda^2)$, with the base point $f_\lambda(p_\lambda)$ if $F(\mathbf{E}_\lambda^2) \not\subset \mathbf{K}^1 - \mathbf{K}^0$, determines an element $R_\lambda \in \mathfrak{F}_k$, where R_λ is the element $1 \in \mathbf{F}_k$ if $F(\mathbf{E}_\lambda^2) \subset \mathbf{K}^1 - \mathbf{K}^0$. Then $\pi_1(\mathbf{K}^n)$ is generated by $a_1, \cdots, a_k$, subject to the relations $R_1 = \cdots = R_{\alpha_2} = 1$. We shall denote this system of generators and relations by R, and shall call it a *natural system,*[23] of generators and relations for $\pi_1(\mathbf{K}^n)$. It is uniquely determined by the choice of V, the choice of orientations for $\mathbf{E}_i^1$ and $\mathbf{E}_\lambda^2$ and of the base points $p_\lambda \in S_\lambda^1$. If $F(\mathbf{E}_\lambda^2) \not\subset \mathbf{K}^1 - \mathbf{K}^0$ let A_λ^2 be a closed simplex in $\mathbf{E}_\lambda^2$, which meets $F(\mathbf{E}_\lambda^2)$ in the base point $f_\lambda(p_\lambda)$ and nowhere else. Let $\tilde{A}_\lambda^2 \subset \tilde{\mathbf{K}}^2$ be the simplex in $u^{-1}(A_\lambda^2)$ which meets $\tilde{\mathbf{V}}$, and let $\tilde{\mathbf{E}}_\lambda^2 \subset \tilde{\mathbf{K}}^2$ be the cell in $u^{-1}(\mathbf{E}_\lambda^2)$ which contains $\tilde{A}_\lambda^2$, or any cell in $u^{-1}(\mathbf{E}_\lambda^2)$ if $F(\mathbf{E}_\lambda^2) \subset \mathbf{K}^1 - \mathbf{K}^0$. Then the chain boundary $\partial\tilde{\mathbf{E}}_\lambda^2$ is given by (3.4) when R_λ is given by (3.3). Therefore the incidence matrix $\mathbf{r}_\lambda^2$ can be calculated algebraically from the system R.

Let $\tilde{\mathbf{E}}_i^q \subset \tilde{\mathbf{K}}^q$ be an arbitrary cell in $u^{-1}(\mathbf{E}_i^q)$ for each $q = 3, \cdots, n$ and $i = 1, \cdots, \alpha_q$. Since the cells $\tilde{\mathbf{E}}_\lambda^{p-1}$ $(\lambda = 1, \cdots, \alpha_{p-1}\,;\, p > 2)$ constitute an homology basis, with coefficients in $\mathfrak{R}$, for the $(p-1)$-cycles in $\tilde{\mathbf{K}}^{p-1}$, and since homology implies equality in the top dimension, it follows that the chain boundaries $\partial\tilde{\mathbf{E}}_i^p$ are given by equations of the form

$$\partial\tilde{\mathbf{E}}_i^p = \sum_{\lambda=1}^{\alpha_{p-1}} r_{i\lambda}^p \tilde{\mathbf{E}}_\lambda^{p-1} \qquad (r_{i\lambda}^p \in \mathfrak{R}).$$

We shall describe the matrices $\mathbf{r}^p = \| r_{i\lambda}^p \|$ $(p = 3, \cdots, n)$, together with the matrices $\mathbf{r}^2$, $\mathbf{r}^1$ which are determined by R, as *natural incidence matrices* for $\mathbf{K}^n$. Since $\mathbf{r}^2$ and $\mathbf{r}^1$ are determined by R, the system $(\mathbf{r}, R)$, which consists of the incidence matrices $\mathbf{r}^n, \cdots, \mathbf{r}^3$, together with R, provides at least as precise, and possibly[24] a more precise description of $\mathbf{K}^n$ than that given by the natural incidence matrices. We shall call $(\mathbf{r}, R)$ a *natural system* for $\mathbf{K}^n$.

We now investigate the conditions under which two complexes may be described by the same natural system. The following arguments lead up to the statement that a natural system for a given complex $\mathbf{K}_0^n$, is unaltered by a

[23] Any system of generators and relations is a natural system for some complex $\mathbf{K}^2$ (cf. O. Veblen, Analysis Situs, New York (1931), chap. V, §24). Thus the system R may be arbitrary except that, for the purposes of this paper, it shall be finite. In particular there may be no generators and $\alpha_2 \geqq 0$ vacuous relations. Or we may have $\alpha_2 = 0$, $k > 0$.

[24] The question implied is: "to what extent is $\mathbf{K}^2$ determined by an abstract knowledge of $\pi_1(\mathbf{K}^2)$, together with given expressions for $\mathbf{r}^2$, $\mathbf{r}^1$, in terms of an arbitrary representation of $\pi_1(\mathbf{K}^2)$?"

'trivial deformation' defined below of $\mathbf{K}_0^n$. Let $\mathbf{K}_0^n$ be a given complex ($n > 0$) and let

$$\mathbf{K}_0^p \to \mathbf{K}_1^p = D_p(\mathbf{K}_0^p) \qquad (0 < p < n)$$

be any formal deformation of $\mathbf{K}_0^p$ into another p-dimensional complex $\mathbf{K}_1^p$. It follows from induction on q ($p < q \leqq n$) and S.S., theorem 14, that the deformation D_p can be extended to a deformation

$$\begin{aligned}\mathbf{K}_0^q \to \mathbf{K}_1^q &= D_q(\mathbf{K}_0^q) \\ &= \mathbf{K}_1^{q-1} + \mathbf{E}_{11}^q + \cdots + \mathbf{E}_{1\alpha_q}^q,\end{aligned}$$

where

$$\mathbf{K}_0^q = \mathbf{K}_0^{q-1} + \mathbf{E}_{01}^q + \cdots + \mathbf{E}_{0\alpha_q}^q$$

and $\mathbf{E}_{1\lambda}^q = D_q(\mathbf{E}_{0\lambda}^q)$. Thus $D_p(\mathbf{K}_0^p)$ can finally be extended to a deformation $D_n(\mathbf{K}_0^n)$, which is also an extension of $D_q(\mathbf{K}_0^q)$ for each $q = p+1, \cdots, n-1$. By S.S., theorem 5, there is a complex $\mathbf{K}_p^*$, which contracts both into $\mathbf{K}_0^p$ and into some subdivision of $\mathbf{K}_1^p$, which, for the purpose of this argument, we may take to be $\mathbf{K}_1^p$ itself. On comparing the proof of S.S., theorem 11, which is given above, with the proof of theorem 14 in S.S., we see that there are complexes $\mathbf{K}_{p+1}^*, \cdots, \mathbf{K}_n^*$, such that

$$\mathbf{K}_q^* = \mathbf{K}_{q-1}^* + C_1^{q+1} + \cdots + C_{\alpha_q}^{q+1} \quad (q = p+1, \cdots, n),$$

where C_λ^{q+1} is a 'deformation cell' for the formal deformation $\mathbf{E}_{0\lambda}^q \to \mathbf{E}_{1\lambda}^q = D_q(\mathbf{E}_{0\lambda}^q)$. The spherical boundary $F(C_\lambda^{q+1})$ consists of $\mathbf{E}_{0\lambda}^q$, $\mathbf{E}_{1\lambda}^q$ and a deformation cylinder for the homotopic deformation $F(\mathbf{E}_{0\lambda}^q) \to F(\mathbf{E}_{1\lambda}^q)$ in $\mathbf{K}_{q-1}^*$. Then $\mathbf{K}_q^*$ contracts into

$$\mathbf{K}_{q-1}^* + \mathbf{K}_i^q = \mathbf{K}_{q-1}^* + \mathbf{E}_{i1}^q + \cdots + \mathbf{E}_{i\alpha_r}^q \qquad (i = 0, 1),$$

and so, by induction on q, into $\mathbf{K}_i^q$. Therefore there is a projection $f_{q-1}(\mathbf{K}_{q-1}^*) = \mathbf{K}_1^{q-1}$ and a projection $g(\mathbf{K}_q^*) = \mathbf{K}_{q-1}^* + \mathbf{K}_1^q$, for a given value of $q > p$, where a projection means a transformation f, such that $f^2 = f$. Let f_{q-1} be extended throughout $\mathbf{K}_{q-1}^* + \mathbf{K}_1^q$ by writing $f_{q-1}(p) = p$ for each point $p \in \mathbf{K}_1^q$. Then $f_q = f_{q-1}g$ is a projection of $\mathbf{K}_q^*$ on $\mathbf{K}_1^q$, such that $f_q(\mathbf{K}_{q-1}^*) = f_{q-1}(\mathbf{K}_{q-1}^*) = \mathbf{K}_1^{q-1}$, and it follows from induction on q that there is a projection $f(\mathbf{K}_n^*) = \mathbf{K}_1^n$, such that $f(\mathbf{K}_q^*) = \mathbf{K}_1^q$ for each $q = p, \cdots, n$. After a suitable deformation of the map f_p we assume that $f(\mathbf{E}_0^0)$ is a vertex $\mathbf{E}_1^0 \in \mathbf{K}_1^0$, where $\mathbf{E}_0^0$ is a given vertex of $\mathbf{K}_0^0$. Let $\tilde{\mathbf{K}}_n^*$ be the universal covering complex of $\tilde{\mathbf{K}}_n^*$, in which a point $\tilde{p} \in \tilde{\mathbf{K}}_n^*$ is defined as a class of curves joining $\mathbf{E}_1^0$ to a given point $p = u(\tilde{p}) \in \mathbf{K}_n^*$. Since $\mathbf{K}_n^*$ contracts into $\mathbf{K}_i^n$ ($i = 0, 1$), whence the latter is a retract by deformation of $\mathbf{K}_n^*$, it follows that $\tilde{\mathbf{K}}_i^n = u^{-1}(\mathbf{K}_i^n)$ is a universal covering complex of $\mathbf{K}_i^n$. Let $\tilde{f}(\tilde{\mathbf{K}}_n^*) = \tilde{\mathbf{K}}_1^n$ be the projection of $\tilde{\mathbf{K}}_n^*$ on $\tilde{\mathbf{K}}_1^n$ in which a given point $\tilde{p} \in \tilde{\mathbf{K}}_n^*$, represented by a segment s, joining $\mathbf{E}_1^0$ to $u(\tilde{p}) \in \mathbf{K}_n^*$, is transformed into the point which is represented by the segment $f(s)$. If $q = 0$, then $\mathbf{K}_q^*$ consists of isolated simply connected complexes, each of which

contracts into a vertex of $\mathbf{K}_i^0$ $(i = 0, 1)$. Similar remarks apply to $\tilde{\mathbf{K}}_0^* = u^{-1}(\tilde{\mathbf{K}}_0^*)$, and it is clear that $\tilde{f}(\tilde{\mathbf{K}}_0^*) = \tilde{\mathbf{K}}_1^0$. In any case the cells in $\tilde{\mathbf{K}}_n^*$, but not in $\mathbf{K}_q^*$, are at least $(q+1)$-dimensional, and hence at least 2-dimensional if $q > 0$. Therefore, if $q > 0$, we may take $s \subset \mathbf{K}_q^*$, whence $f(s) \subset \mathbf{K}_1^q$ and again $\tilde{f}(\tilde{\mathbf{K}}_q^*) = \tilde{\mathbf{K}}_1^q$, where $\tilde{\mathbf{K}}_q^* = u^{-1}(\mathbf{K}_q^*)$. Therefore, $\tilde{f}(\tilde{\mathbf{K}}_q^*) = \tilde{\mathbf{K}}_1^q$ for each $q = p, \cdots, n$ and $p \geqq 0$. Since $\mathbf{K}_n^*$ contracts into $\mathbf{K}_1^n$ it follows that any circuit in $\mathbf{K}_n^*$, which begins and ends at $\mathbf{E}_1^0$, can be deformed into $\mathbf{K}_1^n$, with $\mathbf{E}_1^0$ held fixed.[25] Therefore $\tilde{f}t = t\tilde{f}$, where t is any covering transformation of $\tilde{\mathbf{K}}_n^*$.

Since $\tilde{f}(\tilde{\mathbf{K}}_r^*) = \tilde{\mathbf{K}}_1^r$ $(r = q-1, q; q > p)$, and since homology implies equality between relative q-cycles in $\tilde{\mathbf{K}}_1^q$, mod. $\tilde{\mathbf{K}}_1^{q-1}$, the projection $\tilde{f}$ determines a homomorphism ψ_q, of the group of q-cycles in $\tilde{\mathbf{K}}_q^*$, mod. $\tilde{\mathbf{K}}_{q-1}^*$, on the group of q-cycles in $\tilde{\mathbf{K}}_1^q$, mod. $\tilde{\mathbf{K}}_1^{q-1}$, and $\psi(C^q) = C^q$ if $C^q \subset \tilde{\mathbf{K}}_1^q$. A relative q-cycle in $\tilde{\mathbf{K}}_i^q$, mod. $\tilde{\mathbf{K}}_i^{q-1}$, is simply a q-chain which is composed of the cells in $u^{-1}(\mathbf{E}_{i\lambda}^q)$. Therefore the transformation given by $C^q \to \psi_q(C^q)$, where C^q is an arbitrary q-chain in $\tilde{\mathbf{K}}_0^q$, is a homomorphism of the group of q-chains in $\tilde{\mathbf{K}}_0^q$ in the group of q-chains in $\tilde{\mathbf{K}}_1^q$.

Let R and R' be natural systems of generators and relations for $\pi_1(\mathbf{K}_0^n)$ and $\pi_1(\mathbf{K}_1^n)$, with $\mathbf{E}_0^0$ and $\mathbf{E}_1^0 = f(\mathbf{E}_0^0)$ taken as base points. Let $\pi_1(\mathbf{K}_i^n)$ $(i = 0, 1)$ be identified with the covering group of $\tilde{\mathbf{K}}_i^n$, by taking $\tilde{\mathbf{E}}_i^0$ as a base point in $\tilde{\mathbf{K}}_i^0$, where $\tilde{\mathbf{E}}_i^0$ corresponds to a segment s_i, joining $\mathbf{E}_1^0$ to $\mathbf{E}_i^0$, such that $f(s_i) \subset \mathbf{K}_1^n$ is homotopic to a point.[26] Then the map $f(\mathbf{K}_0^n) = \mathbf{K}_1^n$ determines a natural isomorphism $\psi\{\pi_1(\mathbf{K}_0^n)\} = \pi_1(\mathbf{K}_1^n)$, in which corresponding elements, regarded as covering transformations of $\tilde{\mathbf{K}}_0^n$ and $\tilde{\mathbf{K}}_1^n$, are the transformations induced by the same covering transformation of $\tilde{\mathbf{K}}_n^*$. Let $x \to \bar{x} = \phi(x)$ be the isomorphism of the symbolic group[27] $\mathfrak{G}_R$ on the symbolic group $\mathfrak{G}_{R'}$ $(x \in \mathfrak{G}_R,\ \bar{x} \in \mathfrak{G}_{R'})$, by means of which ψ is expressed algebraically, and let $r \to \bar{r} = \phi(r)$ be the corresponding isomorphism of $\mathfrak{R}$ on $\mathfrak{R}'$, where $\mathfrak{R}$ and $\mathfrak{R}'$ are the group rings of $\mathfrak{G}_R$ and $\mathfrak{G}_{R'}$. Since $\tilde{f}t = t\tilde{f}$, it follows that $\psi_q t = t\psi_q$, whence

$$\psi_q\Big\{\sum_\lambda r_\lambda \tilde{\mathbf{E}}_{0\lambda}^q\Big\} = \sum_\lambda \bar{r}_\lambda \psi_q(\tilde{\mathbf{E}}_{0\lambda}^q), \qquad (q = p+1, \cdots, n), \tag{3.8}$$

where the coefficients r_λ and $\bar{r}_\lambda = \phi(r_\lambda)$ are expressed algebraically as elements of $\mathfrak{R}$ and $\mathfrak{R}'$.

Let $\tilde{\mathbf{E}}_{0\lambda}^q$ be the cell in $u^{-1}(\mathbf{E}_{0\lambda}^q)$, which is chosen as a basis element in defining a set of natural incidence matrices for $\mathbf{K}_0^n$. If $q > p$ we have

$$\tilde{\mathbf{E}}_{0\lambda}^q - \tilde{\mathbf{E}}_{1\lambda}^q = \partial \tilde{C}_\lambda^{q+1} \qquad (\text{mod. } \tilde{\mathbf{K}}_{q-1}^*),$$

[25] See also S.S., theorem 15. This fact has been referred to, by implication, in the statement that $\tilde{\mathbf{K}}_1^n$ is a (single) universal covering complex of $\mathbf{K}_1^n$.

[26] We may take s_0 to be of the form $-f(s) + s$, where s is any segment which joins $\mathbf{E}_1^0$ to $\mathbf{E}_0^0$.

[27] Cf. note III$_5$. $\mathfrak{G}$ stands for the group defined by a given set of generators and relations R.

where $\tilde{C}_\lambda^{q+1}$ and $\tilde{\mathbf{E}}_{1\lambda}^q$ are certain cells in $u^{-1}(C_\lambda^{q+1})$ and $u^{-1}(\mathbf{E}_{1\lambda}^q)$. Since homology implies equality in the top dimensions, and since $\psi_q(C^q) = C^q$ if $C^q \subset \tilde{\mathbf{K}}_1^q$, we have the equality,

$$\psi_q(\tilde{\mathbf{E}}_{0\lambda}^q) = \tilde{\mathbf{E}}_{1\lambda}^q \qquad (\text{mod. } \tilde{\mathbf{K}}_1^{q-1}),$$

between relative q-cycles, which may be interpreted as the absolute equality

$$\psi_q(\tilde{\mathbf{E}}_{0\lambda}^q) = \tilde{\mathbf{E}}_{1\lambda}^q \qquad (q = p+1, \cdots, n), \tag{3.9}$$

between chains. If $q > p+1$ it follows from the relation $\partial\tilde{f} = \tilde{f}\partial$, or $\partial\psi_q = \psi_{q-1}\partial$ and from (3.9) and (3.8), that

$$\begin{aligned} \partial\tilde{\mathbf{E}}_{1\lambda}^q &= \psi_{q-1}(\partial\tilde{\mathbf{E}}_{0\lambda}^q) \\ &= \psi_{q-1}\left(\sum_{a=1}^{\alpha} r_{\lambda a}^q \tilde{\mathbf{E}}_{0a}^{q-1}\right) \qquad (\alpha = \alpha_{q-1}) \\ &= \sum_{a=1}^{\alpha} \bar{r}_{\lambda a}^q \tilde{\mathbf{E}}_{1a}^{q-1} \qquad (q = p+2, \cdots, n). \end{aligned} \tag{3.10}$$

We express by saying that, if $q > p+1$, then the deformation D_n transforms the incidence matrix $\mathbf{r}^q$, for $\mathbf{K}_0^n$ into the incidence matrix $\phi(\mathbf{r}^q) = \| \bar{r}_{\lambda a}^q \|$, for $\mathbf{K}_1^n$.

We now state three conditions under which $\mathfrak{G}_R$ is to be identified with $\mathfrak{G}_{R'}$, and in such a way that ϕ is the identical automorphism. Under these conditions we shall say that $\mathfrak{G}_R$ is unaltered by a deformation $D(\mathbf{K}_0^n) = \mathbf{K}_1^n$. First let $D(\mathbf{K}_0^n)$ be relative to $\mathbf{K}_0^1$, as the term is defined in S.S., p. 255. Then the generators in the system R may be taken as the generators in R', and the two sets of relations will be equivalent (i.e. each set will imply the other). Therefore $\mathfrak{G}_R = \mathfrak{G}_{R'}$. Also both $\mathbf{K}_0^1 = \mathbf{K}_1^1$, and $\mathbf{E}_0^0 = \mathbf{E}_1^0$, whence corresponding elements in the isomorphism $\psi\{\pi_1(\mathbf{K}_0^n)\} = \pi_1(\mathbf{K}_1^n)$ may be represented by the same product of generators of $\mathfrak{G}_R$. Therefore ϕ is the identical automorphism. Secondly let[28] $D = \sigma$, where σ is a stellar sub-division of some simplicial sub-division of $\mathbf{K}^n$. Then, not only shall $\mathfrak{G}_R$ and $\mathfrak{G}_{R'}$ be identified, but also any natural system $(\mathbf{r}, R)$ for $\mathbf{K}^n$ shall be identified with a natural system for $D(\mathbf{K}^n)$. Finally let $(\mathbf{r}, R)$ be a natural system for $\mathbf{K}^n$, and let $\mathbf{V} \subset \mathbf{K}^1$ be the tree used in defining R. After a suitable sub-division of some simplicial sub-division of $\mathbf{K}^n$ (e.g. s_V^2, using the notation explained in S.S., p. 251) let $D(\mathbf{K}^n)$ consist of shrinking an edge in $\mathbf{V}$ into a point.[29] Then D may be copied by a similar deformation of $\tilde{\mathbf{K}}^n$, and $(\mathbf{r}, R)$ obviously determines a natural system for $D(\mathbf{K}^n)$, which we identify with $(\mathbf{r}, R)$. Thus, by reiterating transformations of this form, $\mathbf{V}$ may be shrunk into a point without altering $(\mathbf{r}, R)$.

Let $F(\mathbf{E}_{0\lambda}^p) \to F(\mathbf{E}_{1\lambda}^p)$ $(\lambda = 1, \cdots, \alpha_p\,;\ 2 \leqq p \leqq n)$ be a homotopic deforma-

[28] Cf. S.S., theorem 1.

[29] Cf. S.S., theorem 3.

tion over $\mathbf{K}_0^{p-1}$, where $\mathbf{E}_{01}^p, \cdots, \mathbf{E}_{0\alpha_p}^p$ are the p-cells of a membrane complex $\mathbf{K}_0^n$. Let $\mathbf{K}_1^p = D_p(\mathbf{K}_0^p)$(rel. $\mathbf{K}_0^{p-1}$) be a 'formal extension' of the homotopic deformation $F(\mathbf{E}_{0\lambda}^p) \to F(\mathbf{E}_{1\lambda}^p)$. That is to say

$$\mathbf{K}_1^p = D_p(\mathbf{K}_0^p) \qquad \text{(rel. } \mathbf{K}_0^{p-1}\text{)}$$
$$= \mathbf{K}_0^{p-1} + \mathbf{E}_{11}^p + \cdots + \mathbf{E}_{1\alpha_p}^p,$$

and D_p is given by the construction used in proving S.S., theorem 11, applied to each cell $E_{0\lambda}^p$. We shall call $D_p(\mathbf{K}_0^p)$, or any extension $D_q(\mathbf{K}_0^q)$ $(q = p+1, \cdots, n)$, of $D_p(\mathbf{K}_0^p)$, a *trivial deformation*. If D_p is a trivial deformation the arguments which lead up to (3.9) remain valid when $q = p$. Also $\mathfrak{G}_R = \mathfrak{G}_{R'}$ and $\phi(x) = x$. Therefore (3.10), with $\bar{r}_{\lambda a}^p = r_{\lambda a}^p$, is valid for $q = p+1, \cdots, n$. Since $F(\mathbf{E}_{1\lambda}^p)$ is homotopic to $F(\mathbf{E}_{0\lambda}^p)$ it follows that

$$\partial \tilde{\mathbf{E}}_{1\lambda}^p = \partial \tilde{\mathbf{E}}_{0\lambda}^p,$$

where $\tilde{\mathbf{E}}_{i\lambda}^p$ $(i = 0, 1)$ are the cells which appear in (3.9), with $q = p$. Therefore, with the natural choice of basis cells in $\tilde{\mathbf{K}}_i^p$, the matrix $\mathbf{r}^p$ is unaltered by the deformation. Also $\mathbf{r}^q$ is unaltered if $q < p$, since D_p is relative to $\mathbf{K}_0^{p-1}$. The system R may obviously be taken as a natural system of generators and relations for $\pi_1\{D(\mathbf{K}_0^n)\}$, provided we choose a suitable base point on $F(\mathbf{E}_{1\lambda}^2)$ if $p = 2$. Therefore any natural system, $(\mathbf{r}, R)$, for $\mathbf{K}_0^n$, determines a natural system for $D_n(\mathbf{K}_0^n)$, which we identify with $(\mathbf{r}, R)$. We express this by saying that a *natural system*, $(\mathbf{r}, R)$, *for* $\mathbf{K}_0^n$, *is unaltered by a trivial deformation.*

We conclude this section with a theorem which, together with theorems 8 and 9 below, is a measure of the effectiveness of the theory developed here. The theorem may be compared with note III$_3$, in §6 below.

THEOREM 5. *If $n = 3$, or if $n > 3$ and $\pi_s(\mathbf{K}^n) = 0$ for $s = 2, \cdots, n-2$, then the nucleus of membrane complex $\mathbf{K}^n$ is completely determined by one of its natural systems.*

Let $(\mathbf{r}, R)$ be a natural system both for $\mathbf{K}_0^n$ and for $\mathbf{K}_1^n$, and if $n > 3$ let $\pi_s(\mathbf{K}_\rho^n) = 0$ for $\rho = 0, 1$; $s = 2, \cdots, n-2$. After shrinking a tree in $\mathbf{K}_\rho^1$ into a point, we assume that $\mathbf{K}_\rho^0$ is a single vertex $\mathbf{E}_\rho^0$, and $\mathbf{K}_\rho^1$ a set of circuits which begin and end at $\mathbf{E}_\rho^0$. Since the system R is the same for both complexes it determines a $(1-1)$ correspondence between the circuits in $\mathbf{K}_0^1$ and in $\mathbf{K}_1^1$. We identify each circuit in $\mathbf{K}_0^1$ with its image in this correspondence. Since the relations $R_\lambda = 1$ are the same for both complexes we have $\mathbf{K}_1^2 = D_2(\mathbf{K}_0^2)$, where D_2 is a trivial deformation, which we extend throughout $\mathbf{K}_0^n$. Therefore we assume that $\mathbf{K}_0^2 = \mathbf{K}_1^2$. Let $\mathbf{K}_0^p = \mathbf{K}_1^p = \mathbf{K}^p$, say, where $2 \leqq p < n$. Since the matrix $\mathbf{r}^{p+1}$ is the same for $\mathbf{K}_0^{p+1}$ and for $\mathbf{K}_1^{p+1}$ it follows that

$$\partial \tilde{\mathbf{E}}_{0i}^{p+1} \sim \partial \tilde{\mathbf{E}}_{1i}^{p+1} \qquad \text{in} \qquad \tilde{\mathbf{K}}^p,$$

where $\mathbf{E}_{\rho i}^{p+1}$ $(i = 1, \cdots, \alpha_{p+1})$ are the $(p+1)$-cells of $\mathbf{K}_\rho^{p+1}$ and $\tilde{\mathbf{E}}_{\rho i}^{p+1} \subset \tilde{\mathbf{K}}_\rho^{p+1}$ $(\rho = 0, 1)$ is the chosen basis element which covers $\mathbf{E}_{\rho i}^{p+1}$. Since $\pi_s(\mathbf{K}^p) = \pi_s(\mathbf{K}_\rho^n) = 0$ for $s = 2, \cdots, p-1$, and since $\pi_s(\mathbf{K}^p) \simeq \pi_s(\tilde{\mathbf{K}}^p)$ if $s > 1$, where $\simeq$ denotes isomorphism, we have $\pi_s(\tilde{\mathbf{K}}^p) = 0$ for $s = 1, \cdots, p-1$. Therefore

two maps $f_0(S^p), f_1(S^p) \subset \tilde{\mathbf{K}}^p$ are homotopic to each other if[30] $f_0(S^p) \sim f_1(S^p)$. Therefore $F(\tilde{\mathbf{E}}_{0i}^{n+1})$ is homotopic in $\tilde{\mathbf{K}}^p$ to $F(\tilde{\mathbf{E}}_{1i}^{p+1})$ for each $i = 1, \cdots, \alpha_{p+1}$. Therefore $F(\mathbf{E}_{0i}^{p+1})$ is homotopic in $\mathbf{K}^p$ to $F(\mathbf{E}_{1i}^{p+1})$, whence $\mathbf{K}_1^{p+1} = D_{p+1}(\mathbf{K}_0^{p+1})$, where D_{p+1} is a trivial deformation. The deformation D_{p+1} may be extended throughout $\mathbf{K}_0^n$, and the theorem follows from induction on p.

Notice that, if $\mathbf{K}^1 = \mathbf{K}^0$ or if $\mathbf{K}^1$ is a single circuit, then theorem 5 remains valid if $(\mathbf{r}, R)$ is replaced by a set of natural incidence matrices. For in this case R is determined by $\mathbf{r}^2$ and $\mathbf{r}^1$, and even by the first and second integral incidence matrices.

4. We now consider a system $(\mathbf{r}, R)$ from a purely algebraic point of view. We take R to be a particular (finite) system of generators and relations for a symbolic group $\mathfrak{G}_R$, and $\mathbf{r} = (\mathbf{r}^n, \cdots, \mathbf{r}^3)$ to be a set of matrices whose elements are in the group ring $\mathfrak{R}$, of $\mathfrak{G}_R$. The matrix $\mathbf{r}^{p+1}$ is to have α_{p+1} rows and α_p columns $(p = 2, \cdots, n-1; \alpha_p \geqq 0)$, where α_2 is the number of relations in the system R, which may be empty in the sense that it has α_2 relations and no generators. Let $\mathbf{r}^2$ and $\mathbf{r}^1$ be the matrices which are calculated from R by the methods explained in §3. Then it follows from (3.7) that $\mathbf{r}^2\mathbf{r}^1 = 0$, and we require that $\mathbf{r}^p\mathbf{r}^{p-1} = 0$ for $p = 3, \cdots, n$. We shall call $\mathbf{r}^n, \cdots, \mathbf{r}^1$ the *incidence matrices* associated with the system $(\mathbf{r}, R)$. We classify such systems by definitions of L- and L^*-equivalence. These are stated in terms of certain elementary transformations of R, together with transformations of the kind considered in §2, which are applicable to the matrices $\mathbf{r}$. These two sets of transformations are not independent, since the transformations of $\mathbf{r}^3$ from the right are governed by the transformations of R, and conversely. This interlocking between the two sets of transformations enables us to avoid a familiar difficulty, which arises in dealing with deformations in a 2-dimensional complex and is absent in the higher dimensionalities.[31]

The elementary transformations of R shall consist of the following transformations:

T_1. *adding a new generator a_0, and a new relation of the form $xa_0^\epsilon y = 1$ $(\epsilon = \pm 1)$, where x and y are products of the original generators;*[32]

T_2. *the inverse of T_1;*

T_3. *adding a new relation which is a consequence of the original relations;*

T_4. *the inverse of T_3.*

It is to be understood any generator a_i, together with a relation of the form $xa_i^\epsilon y = 1$, may be discarded by a transformation T_2, provided $a_i^{\pm 1}$ does not occur in x, y or in any of the other relations. If $R' = T_1(R)$ the symbolic group $\mathfrak{G}_{R'}$ is to be regarded as distinct from $\mathfrak{G}_R$, and we associate with T_1 the isomorphism $\phi(\mathfrak{G}_R) = \mathfrak{G}_{R'}$, such that $\phi(a_i) = a_i$, where $a_1, \cdots, a_k$ are the generators of R.

[30] W. Hurewicz, Proc. Kon. Akad. Amsterdam, 38 (1935), 521–8, theorem 1, p. 522.

[31] See, for example, theorem VII in H. Hopf and E. Pannwitz, Math. Annalen, 108 (1933), 433–65.

[32] By a product of the generators we mean a product of the generators and their inverses.

With the transformation $R' = T_2(R)$ we associate the isomorphism $\phi(G_R) = G_{R'}$, such that $\phi(a_i) = a_i$, $\phi(a_0) = (yx)^{-\epsilon}$, where a_0 is the generator and $xa_0^{\epsilon}y = 1$ is the relation to be discarded. If $R' = T_i(R)$, where $i = 3$ or 4, then $\mathfrak{G}_R = \mathfrak{G}_{R'}$ and ϕ shall be the identical automorphism of $\mathfrak{G}_R$. Thus there is, in each case, a definite isomorphism $\phi(\mathfrak{G}_R) = \mathfrak{G}_{R'}$ associated with the transformation T_i. Therefore a given transformation $T(R) = R'$, where T is the resultant of transformations of the form $T_1, \cdots, T_4$, determines a unique isomorphism $\phi(\mathfrak{G}_R) = \mathfrak{G}_{R'}$. It follows from the standard proof of a familiar theorem[33] that any isomorphism of $\mathfrak{G}_R$ on an isomorphic group, which is given by a finite system of generators and relations, can be expressed as a product of the transformations $T_1, \cdots, T_4$. But, as we shall see, a system R, as part of a system $(\mathbf{r}, R)$, cannot necessarily be transformed into an equivalent system $R' \subset (\mathbf{r}', R')$, since T_4 is to be forbidden except when it is consistent with a corresponding transformation of the matrix $\mathbf{r}^3$.

We now consider the transformations T_i in relation to the matrices $\mathbf{r}^n, \cdots, \mathbf{r}^3$. Let $R' = T_i(R)$ and let $\phi(\mathfrak{R}) = \mathfrak{R}'$ be the isomorphism between the group rings of $\mathfrak{G}_R$ and $\mathfrak{G}_{R'}$, which is determined by the corresponding isomorphism $\phi(\mathfrak{G}_R) = \mathfrak{G}_{R'}$. We first of all replace $\mathbf{r}^p$ by the matrix[34] $\phi(\mathbf{r}^p)$ $(p = 3, \cdots, n)$. Then T_1 shall be accompanied by the addition of a 0^{th} column of zeros to $\phi(\mathbf{r}^3)$. If T_2 is applicable to R, then the i^{th} column of $\mathbf{r}^2$ has $\pm x$ in the λ^{th} row, for some $x \in \mathfrak{G}_R$, and zeros everywhere else, where a_i is the generator and $R_\lambda = 1$ the relation to be discarded. It follows from the relation $\mathbf{r}^3\mathbf{r}^2 = 0$ that the λ^{th} column of $\phi(\mathbf{r}^3)$ consists entirely of zeros, and T_2 is to be accompanied by the removal of this column from $\phi(\mathbf{r}^3)$. Let $R_0 = 1$ be the relation added by T_3, where

$$R_0 = x_1 R_{\lambda_1}^{\epsilon_1} x_1^{-1} \cdots x_q R_{\lambda_q}^{\epsilon_q} x_q^{-1} \qquad (x_i \in \mathfrak{G}_R). \tag{4.1}$$

Then

$$\partial \tilde{\mathbf{E}}_0^2 = \epsilon_1 x_1 \partial \tilde{\mathbf{E}}_{\lambda_1}^2 + \cdots + \epsilon_q x_q \partial \tilde{\mathbf{E}}_{\lambda_q}^2, \tag{4.2}$$

or, after simplification,

$$\partial \tilde{\mathbf{E}}_0^2 = c_1 \partial \tilde{\mathbf{E}}_1^2 + \cdots + c_{\alpha_2} \partial \tilde{\mathbf{E}}_{\alpha_2}^2, \tag{4.3}$$

for certain values of $c_i \in \mathfrak{R}$, where $\partial \tilde{\mathbf{E}}_t^2$ $(t = 0, \cdots, \alpha_2)$ is the linear form (3.4) when R_t is the x in (3.3), and $\tilde{\mathbf{E}}_i^1$ are now to be regarded as undefined basis elements for a modulus with coefficients in $\mathfrak{R}$. The addition of $R_0 = 1$ to the

[33] K. Reidemeister, Einführung in die kombinatorische Topologie, Brunswick (1932), 48.

[34] If $T = T_1$ the same formal expressions may be used to represent $r_{\lambda a}^p$ as $\phi(r_{\lambda a}^p)$, only $r_{\lambda a}^p$ and $\phi(r_{\lambda a}^p)$ must be interpreted as elements of different rings. If $T = T_2$ we first of all substitute yx for $a_0^{-\epsilon}$, whenever the latter appears in given formal expressions for $r_{\lambda a}^p$. This does not alter the element $r_{\lambda a}^p \in \mathfrak{R}$, only its formal expression as a polynomial in the generators. The transformation $r_{\lambda a}^p \to \phi(r_{\lambda a}^p)$ may then be defined as when $T = T_1$. If $T = T_3$ or T_4 we have $\phi(\mathbf{r}^p) = \mathbf{r}^p$.

relations of R is to be accompanied by an expanion of the form

$$\mathbf{r}^3 \to \left\| \begin{matrix} \epsilon x & -\epsilon x \mathbf{c} \\ 0 & \mathbf{r}^3 \end{matrix} \right\|, \tag{4.4}$$

where $\epsilon = \pm 1$, $x \in \mathfrak{G}_R$ is arbitrary and $-\epsilon x \mathbf{c}$ stands for the row $\| -\epsilon x c_1, \cdots, -\epsilon x c_{\alpha_2} \|$ together with the corresponding secondary expansion of $\mathbf{r}^4$. Conversely, let (4.4) be a given expansion, in which x and c_i, and hence $-\epsilon x c_i$, are arbitrary, and let (4.3) be expressed in the form (4.2). Then we accompany the expansion (4.4) by adding the new relation $R_0 = 1$, where R_0 is of the form (4.1), with the factors $x_i R_{\lambda_i}^{\epsilon_i} x_i^{-1}$ arranged in an arbitrary order. A transformation of the form T_4, in which $R_\lambda = 1$ is the relation to be discarded, shall only be applied if the λ^{th} column of $\mathbf{r}^3$ has $\pm x$ ($x \in \mathfrak{G}_R$) in the i^{th} row, for an arbitrary $i = 1, \cdots, \alpha_3$, and zeros everywhere else. In this case it is to imply the contraction of $\mathbf{r}^3$, in which the i^{th} row and the λ^{th} column are discarded, and the corresponding contraction of $\mathbf{r}^4$. Conversely, such a contraction of $\mathbf{r}^3$ is only to be applied when $R_\lambda = 1$ is a consequence of the other relations. In this case it is to be accompanied by the transformation T_4, in which the relation $R_\lambda = 1$ is discarded. It is now to be understood that a transformation $T_i(\mathbf{r}, R)$ consists of the transformation $R \to T_i(R)$, together with the transformation $\mathbf{r}^p \to T_i(\mathbf{r}^p)$, where $T_i(\mathbf{r}^4)$, $T_i(\mathbf{r}^3)$ are the matrices into which $\mathbf{r}^4$, $\mathbf{r}^3$ are transformed by the above rules, and $T_i(\mathbf{r}^p) = \phi(\mathbf{r}^p)$ if $p > 4$, or if $p = 4$ and $i = 1$ or 2.

In addition to the transformations T_i we allow the matrices $\mathbf{r}^n, \cdots, \mathbf{r}^4$ to be transformed as in §2. That is to say, we allow arbitrary expansions and contractions of $\mathbf{r}^p$ ($p = 4, \cdots, n$), which must be accompanied by the appropriate secondary expansions and contractions of $\mathbf{r}^3$ if $p = 4$, and transformations of the form

$$\mathbf{r}^{p+1} \to \mathbf{r}^{p+1}\mathbf{a}_p^{-1}, \qquad \mathbf{r}^p \to \mathbf{a}_p \mathbf{r}^p \qquad (p = 3, \cdots, n),$$

where $\mathbf{a}_p$ is any regular, or elementary, matrix of degree α_p, according to the kind of equivalence considered. Two systems $(\mathbf{r}, R)$ and $(\mathbf{r}', R')$, will be described as *X-equivalent* ($X = L$ or L^*) if, and only if, they are equivalent under the combined set of transformations, in which $\mathbf{a}_p$ may be any regular matrix, of the appropriate degree, if $X = L^*$, and is to be an elementary matrix if $X = L$. More precisely, we shall describe $(\mathbf{r}, R)$ and $(\mathbf{r}', R')$ as *X-equivalent under the isomorphism* ϕ, where $\phi(\mathfrak{G}_R) = \mathfrak{G}_{R'}$ is the isomorphism induced by the transformation $T(R) = R'$. It is important to notice that ϕ depends on the actual sequence of the transformations T_i which appear in the transformation T, not only on the final result. For example, ϕ need not be the identical automorphism if $\mathfrak{G}_R = \mathfrak{G}_{R'}$, or even if $R' = R$. Clearly theorems 1 and 2 are valid when restated in terms of equivalence between systems $(\mathbf{r}, R)$ and $(\mathbf{r}', R')$. Let us describe as an *elementary expansion* of $(\mathbf{r}, R)$ either a transformation of the form T_1, or T_3, or an expansion of the form (2.5), where $p > 3$, and let us describe the inverse of an elementary expansion as an *elementary contraction*. It follows from the proof of lemma 4 that $(\mathbf{r}, R)$ can be transformed

into any L-equivalent system by a series of elementary expansions and contractions, provided we allow primary expansions of the empty matrix $\mathbf{r}^{n+1}$.

Let κ_ν denote the ν^{th} column of $\mathbf{r}^3$, let x be a given element of $\mathfrak{G}_R$, and let λ, μ $(\lambda \neq \mu)$ be particular values of $\nu = 1, \cdots, \alpha_2$. I say that the transformation

$$R_\lambda \to R'_\lambda = xR_\lambda^\epsilon x^{-1} R_\mu \qquad (\epsilon = \pm 1), \tag{4.5}$$

accompanied by $\kappa_\lambda \to \epsilon\kappa_\lambda x^{-1}$, $\kappa_\mu \to \kappa_\mu - \epsilon\kappa_\lambda x^{-1}$, is the resultant of a transformation T_3, followed by elementary transformations of $\mathbf{r}^4$, $\mathbf{r}^3$, whose resultant leaves $\mathbf{r}^4$ unaltered, followed by T_4. For we first add the relation $R'_\lambda = 1 (R'_\lambda = R_0)$ to R and the row r^3_{0i} $(i = 0, \cdots, \alpha_2)$ to $\mathbf{r}^3$, where

$$r^3_{00} = 1, \qquad r^3_{0\lambda} = -\epsilon x, \qquad r^3_{0\mu} = -1, \qquad r^3_{0\nu} = 0 \qquad (\nu \neq \lambda, \mu),$$

and the columns $r^3_{a0} = \delta_{a0}$, $r^4_{k0} = 0$ $(a = 0, \cdots, \alpha_3 ; k = 1, \cdots, \alpha_4)$ to $\mathbf{r}^3$ and $\mathbf{r}^4$. We then clear away the elements of $\mathbf{r}^3$, which stand under $r^3_{0\lambda}$, by the transformations $\rho_t \to \rho_t + \epsilon r^3_{t\lambda} x^{-1} \rho_0$ $(t = 1, \cdots, \alpha_3)$, where ρ_t stands for the t^{th} row of $\mathbf{r}^3$. As in lemma 4, it follows from the relation $\mathbf{r}^4\mathbf{r}^3 = 0$ that the corresponding transformation of $\| 0, \mathbf{r}^4 \|$ leaves the latter unchanged. We now discard the relation $R_\lambda = 1$, and the appropriate rows and columns of the expanded matrices $\mathbf{r}^4$, $\mathbf{r}^3$, and rearrange the notation so that $R'_\lambda = 1$ appears as the λ^{th} relation. The final result is the transformation (4.5). Notice that the transformation $R_\lambda \to xR_\lambda^\epsilon x^{-1}$ is the resultant of transformations of the form (4.5), namely

$$R_\lambda \to xR_\lambda^{-\epsilon} x^{-1} R_\mu \to R_\mu (xR_\lambda^{-\epsilon} x^{-1} R_\mu)^{-1} R_\mu^{-1} R_\mu = xR_\lambda^\epsilon x^{-1}.$$

Let $(\mathbf{r}, R)$ and $(\mathbf{r}', R')$ be two given systems, and let $\mathbf{r}^2$, $\mathbf{r}^1$ and $\mathbf{r}'^2$, $\mathbf{r}'^1$ be the second and first two incidence matrices determined by R and R'.

THEOREM 6. *If the systems* $(\mathbf{r}, R)$ *and* $(\mathbf{r}', R')$ *are* X*-equivalent* $(X = L$ *or* $L^*)$ *under a given isomorphism* ϕ*, then the sets of incidence matrices* $\mathbf{r}^n, \cdots, \mathbf{r}^1$ *and* $\mathbf{r}'^n, \cdots, \mathbf{r}'^1$ *are* X*-equivalent under* ϕ.

This relation between $\mathbf{r}^n, \cdots, \mathbf{r}^1$ and $\mathbf{r}'^n, \cdots, \mathbf{r}'^1$ is obviously transitive, and if $R' = R$ and R is left unaltered throughout the passage from $(\mathbf{r}, R)$ to $(\mathbf{r}', R')$ there is nothing to prove. Therefore the theorem will follow from an inductive argument when we have proved it in case T is a single transformation T_i. If $i = 3$ or 4 the theorem is implicit in the definition of $T_i(\mathbf{r}, R)$. On replacing ϕ by ϕ^{-1} the equivalence relation between $\mathbf{r}^n, \cdots, \mathbf{r}^1$ and $\mathbf{r}'^n, \cdots, \mathbf{r}'^1$ is seen to be symmetric. Therefore it is sufficient to consider the case where $i = 1$. If a relation $R_0 = xa_0^\epsilon y = 1$ is replaced by $zR_0^{\pm 1} z^{-1} = 1$ $(z \in \mathfrak{G}_R)$, the 0^{th} row of $\phi(\mathbf{r}^2)$ is multiplied by $\pm z$. Therefore we may assume that $R_0 = 0_0 x$, for some x which is a product of the existing generators. Then the 0^{th} row of $\mathbf{r}'^2$ is $\| 1, a_0 r_1, \cdots, a_0 r_k \|$, where r_i means the same as it does in (3.4) when x is given by (3.3). The i^{th} row of $\mathbf{r}'^1$ is $a_i - 1$ $(i = 0, \cdots, k)$, and since $x = a_0^{-1}$, in consequence of $R_0 = 1$, it follows from (3.7) that

$$\begin{aligned} a_0 - 1 &= a_0(1 - x) \\ &= -\sum_{i=1}^{k} a_0 r_i (a_i - 1), \end{aligned}$$

whence $\mathbf{r}'^2$, $\mathbf{r}'^1$ are obtained from $\phi(\mathbf{r}^2)$, $\phi(\mathbf{r}^1)$ by an expansion of the form (2.5). The appropriate column of zeros is added to $\phi(\mathbf{r}^3)$, and the theorem is established.

We now prove a theorem which shows how the generators of R can be manipulated by what may be called the 'free group calculus.' Let the transformation

$$a_i \to f_i(a_1', \cdots, a_k') = f_i(a') \qquad (i = 1, \cdots, k)$$

determine an isomorphism of $\mathfrak{F}_k$ on the free group $\mathfrak{F}_k'$, which is freely generated by $a_1', \cdots, a_k'$. Some or all of the generators $a_1', \cdots, a_k'$ may coincide with some of $a_1^{\pm 1}, \cdots, a_k^{\pm 1}$, so that the possibility $\mathfrak{F}_k' = \mathfrak{F}_k$ is not excluded. Let

$$R_\lambda' = R_\lambda\{f_1(a'), \cdots, f_k(a')\} \qquad (\lambda = 1, \cdots, \alpha_2),$$

where $R_\lambda(a_1, \cdots, a_k) = 1$ are the relations of the system R, and let R' be the system which consists of the generators $a_1', \cdots, a_k'$, subject to the relations $R_\lambda' = 1$. Since the transformation given by $a_i \to f_i(a')$ is an isomorphism of $\mathfrak{F}_k$ on $\mathfrak{F}_k'$, the group $\mathfrak{G}_{R'}$ is also generated by $f_1(a'), \cdots, f_k(a')$, and $a_i \to f_i(a')$ determines an isomorphism $\phi(\mathfrak{G}_R) = \mathfrak{G}_{R'}$, in which $\phi(a_i) = f_i(a')$. Let $\mathbf{r}'^s = \phi(\mathbf{r}^s)$ $(s = 3, \cdots, n)$, where $\mathbf{r}^n, \cdots, \mathbf{r}^3$ are the matrices of a given system $(\mathbf{r}, R)$. Then we have the theorem:

THEOREM 7. *The system* $(\mathbf{r}, R)$ *is L-equivalent under the isomorphism* ϕ *to the system* $(\mathbf{r}', R')$.

First assume that $a_i' \neq a_j^{\pm 1}$ for any values of i, j and that either

1. $f_p(a') = a_p'^{\epsilon}$, $\quad f_i(a') = a_i'(\epsilon = \pm 1, i \neq p)$, or
2. $f_p(a') = a_p'a_q'$, $f_i(a') = a_i'$ $(q, i \neq p)$.

Let $a_i' \to g_i(a)$ be the inverse of the isomorphism $a_i \to f_i(a')$, so that $g_p(a) = a_p^{\epsilon}$, $g_i(a) = a_i$, in the first case, and $g_p(a) = a_p a_q^{-1}$, $g_i(a) = a_i$ in the second. Let $T(R) = \bar{R}$ be the expansion of R in which $a_1', \cdots, a_k'$ are added as new generators, together with the relations $a_i'^{-1}g_i(a) = 1$, let $\phi_1(\mathfrak{G}_R) = \mathfrak{G}_{\bar{R}}$, given by $\phi_1(a_i) = a_i$, be the corresponding isomorphism of $\mathfrak{G}_R$, and let $\bar{\mathbf{r}}^3 = T\{\phi_1(\mathbf{r}^3)\}$. Then $\bar{\mathbf{r}}^3 = \| \phi_1(\mathbf{r}^3), 0 \|$, where 0 indicates the columns which correspond to the new relations. The latter can obviously be transformed into the relations $a_i^{-1}f_i(a') = 1$ by transformations of the form (4.5), and these transformations leave $\bar{\mathbf{r}}^3$ unaltered, since the new columns consist entirely of zeros. We now transform $R_\lambda(a_1, \cdots, a_k) = 1$ into $R_\lambda'(a_1', \cdots, a_k') = 1$, for each $\lambda = 1, \cdots, \alpha_2$, by substituting $f_i(a')$ for a_i throughout the products R_λ. This operation is the resultant of transformations of the form

$$xa_iy' \to xa_iy'y'^{-1}a_i^{-1}f_iy' \qquad \text{or} \qquad xa_i^{-1}y' \to xf_i^{-1}a_ix^{-1}xa_i^{-1}y',$$

where $f_i = f_i(a')$, $xa_i^{\epsilon}y = R_\lambda$, for some $\lambda = 1, \cdots, \alpha_2$, and y' is the result of substituting $f_i(a')$ for a_i throughout the product y. Therefore it follows without difficulty that the transformation $R_\lambda \to R_\lambda'$ is the resultant of transformations of the form (4.5). Moreover the columns of $\mathbf{r}^3$, which correspond to the original relations $R_\lambda = 1$, are unaltered by these transformations.[35] Finally we discard

[35] As explained above, in the definition of $T_2(\mathbf{r}^3)$, it follows from the condition $\mathbf{r}^3\mathbf{r}^2 = 0$, and the form of the relations $a_i^{-1}f_i(a') = 1$, that the remaining (zero) columns also are ultimately left unaltered.

into any L-equivalent system by a series of elementary expansions and contractions, provided we allow primary expansions of the empty matrix $\mathbf{r}^{n+1}$.

Let κ_ν denote the ν^{th} column of $\mathbf{r}^3$, let x be a given element of $\mathfrak{G}_R$, and let λ, μ ($\lambda \neq \mu$) be particular values of $\nu = 1, \cdots, \alpha_2$. I say that the transformation

$$R_\lambda \to R'_\lambda = xR_\lambda^\epsilon x^{-1} R_\mu \qquad (\epsilon = \pm 1), \tag{4.5}$$

accompanied by $\kappa_\lambda \to \epsilon\kappa_\lambda x^{-1}$, $\kappa_\mu \to \kappa_\mu - \epsilon\kappa_\lambda x^{-1}$, is the resultant of a transformation T_3, followed by elementary transformations of $\mathbf{r}^4$, $\mathbf{r}^3$, whose resultant leaves $\mathbf{r}^4$ unaltered, followed by T_4. For we first add the relation $R'_\lambda = 1 (R'_\lambda = R_0)$ to R and the row r^3_{0i} ($i = 0, \cdots, \alpha_2$) to $\mathbf{r}^3$, where

$$r^3_{00} = 1, \qquad r^3_{0\lambda} = -\epsilon x, \qquad r^3_{0\mu} = -1, \qquad r^3_{0\nu} = 0 \qquad (\nu \neq \lambda, \mu),$$

and the columns $r^3_{a0} = \delta_{a0}$, $r^4_{k0} = 0$ ($a = 0, \cdots, \alpha_3$; $k = 1, \cdots, \alpha_4$) to $\mathbf{r}^3$ and $\mathbf{r}^4$. We then clear away the elements of $\mathbf{r}^3$, which stand under $r^3_{0\lambda}$, by the transformations $\rho_t \to \rho_t + \epsilon r^3_{t\lambda} x^{-1} \rho_0$ ($t = 1, \cdots, \alpha_3$), where ρ_t stands for the t^{th} row of $\mathbf{r}^3$. As in lemma 4, it follows from the relation $\mathbf{r}^4\mathbf{r}^3 = 0$ that the corresponding transformation of $\| 0, \mathbf{r}^4 \|$ leaves the latter unchanged. We now discard the relation $R_\lambda = 1$, and the appropriate rows and columns of the expanded matrices $\mathbf{r}^4$, $\mathbf{r}^3$, and rearrange the notation so that $R'_\lambda = 1$ appears as the λ^{th} relation. The final result is the transformation (4.5). Notice that the transformation $R_\lambda \to xR_\lambda^\epsilon x^{-1}$ is the resultant of transformations of the form (4.5), namely

$$R_\lambda \to xR_\lambda^{-\epsilon} x^{-1} R_\mu \to R_\mu (xR_\lambda^{-\epsilon} x^{-1} R_\mu)^{-1} R_\mu^{-1} R_\mu = xR_\lambda^\epsilon x^{-1}.$$

Let $(\mathbf{r}, R)$ and $(\mathbf{r}', R')$ be two given systems, and let $\mathbf{r}^2$, $\mathbf{r}^1$ and $\mathbf{r}'^2$, $\mathbf{r}'^1$ be the second and first two incidence matrices determined by R and R'.

THEOREM 6. *If the systems* $(\mathbf{r}, R)$ *and* $(\mathbf{r}', R')$ *are* X*-equivalent* ($X = L$ *or* L^*) *under a given isomorphism* ϕ, *then the sets of incidence matrices* $\mathbf{r}^n, \cdots, \mathbf{r}^1$ *and* $\mathbf{r}'^n, \cdots, \mathbf{r}'^1$ *are* X*-equivalent under* ϕ.

This relation between $\mathbf{r}^n, \cdots, \mathbf{r}^1$ and $\mathbf{r}'^n, \cdots, \mathbf{r}'^1$ is obviously transitive, and if $R' = R$ and R is left unaltered throughout the passage from $(\mathbf{r}, R)$ to $(\mathbf{r}', R')$ there is nothing to prove. Therefore the theorem will follow from an inductive argument when we have proved it in case T is a single transformation T_i. If $i = 3$ or 4 the theorem is implicit in the definition of $T_i(\mathbf{r}, R)$. On replacing ϕ by ϕ^{-1} the equivalence relation between $\mathbf{r}^n, \cdots, \mathbf{r}^1$ and $\mathbf{r}'^n, \cdots, \mathbf{r}'^1$ is seen to be symmetric. Therefore it is sufficient to consider the case where $i = 1$. If a relation $R_0 = xa_0^\epsilon y = 1$ is replaced by $zR_0^{\pm 1} z^{-1} = 1$ ($z \in \mathfrak{G}_R$), the 0^{th} row of $\phi(\mathbf{r}^2)$ is multiplied by $\pm z$. Therefore we may assume that $R_0 = 0_0 x$, for some x which is a product of the existing generators. Then the 0^{th} row of $\mathbf{r}'^2$ is $\| 1, a_0 r_1, \cdots, a_0 r_k \|$, where r_i means the same as it does in (3.4) when x is given by (3.3). The i^{th} row of $\mathbf{r}'^1$ is $a_i - 1$ ($i = 0, \cdots, k$), and since $x = a_0^{-1}$, in consequence of $R_0 = 1$, it follows from (3.7) that

$$\begin{aligned} a_0 - 1 &= a_0(1 - x) \\ &= -\sum_{i=1}^{k} a_0 r_i (a_i - 1), \end{aligned}$$

whence $\mathbf{r}'^2$, $\mathbf{r}'^1$ are obtained from $\phi(\mathbf{r}^2)$, $\phi(\mathbf{r}^1)$ by an expansion of the form (2.5). The appropriate column of zeros is added to $\phi(\mathbf{r}^3)$, and the theorem is established.

We now prove a theorem which shows how the generators of R can be manipulated by what may be called the 'free group calculus.' Let the transformation

$$a_i \to f_i(a_1', \cdots, a_k') = f_i(a') \qquad (i = 1, \cdots, k)$$

determine an isomorphism of $\mathfrak{F}_k$ on the free group $\mathfrak{F}_k'$, which is freely generated by $a_1', \cdots, a_k'$. Some or all of the generators $a_1', \cdots, a_k'$ may coincide with some of $a_1^{\pm 1}, \cdots, a_k^{\pm 1}$, so that the possibility $\mathfrak{F}_k' = \mathfrak{F}_k$ is not excluded. Let

$$R_\lambda' = R_\lambda\{f_1(a'), \cdots, f_k(a')\} \qquad (\lambda = 1, \cdots, \alpha_2),$$

where $R_\lambda(a_1, \cdots, a_k) = 1$ are the relations of the system R, and let R' be the system which consists of the generators $a_1', \cdots, a_k'$, subject to the relations $R_\lambda' = 1$. Since the transformation given by $a_i \to f_i(a')$ is an isomorphism of $\mathfrak{F}_k$ on $\mathfrak{F}_k'$, the group $\mathfrak{G}_{R'}$ is also generated by $f_1(a'), \cdots, f_k(a')$, and $a_i \to f_i(a')$ determines an isomorphism $\phi(\mathfrak{G}_R) = \mathfrak{G}_{R'}$, in which $\phi(a_i) = f_i(a')$. Let $\mathbf{r}'^s = \phi(\mathbf{r}^s)$ $(s = 3, \cdots, n)$, where $\mathbf{r}^n, \cdots, \mathbf{r}^3$ are the matrices of a given system $(\mathbf{r}, R)$. Then we have the theorem:

THEOREM 7. *The system* $(\mathbf{r}, R)$ *is L-equivalent under the isomorphism* ϕ *to the system* $(\mathbf{r}', R')$.

First assume that $a_i' \neq a_j^{\pm 1}$ for any values of i, j and that either

1. $f_p(a') = a_p'^{\epsilon}$, $\quad f_i(a') = a_i'(\epsilon = \pm 1, i \neq p)$, or
2. $f_p(a') = a_p' a_q'$, $f_i(a') = a_i'$ $(q, i \neq p)$.

Let $a_i' \to g_i(a)$ be the inverse of the isomorphism $a_i \to f_i(a')$, so that $g_p(a) = a_p^{\epsilon}$, $g_i(a) = a_i$, in the first case, and $g_p(a) = a_p a_q^{-1}$, $g_i(a) = a_i$ in the second. Let $T(R) = \bar{R}$ be the expansion of R in which $a_1', \cdots, a_k'$ are added as new generators, together with the relations $a_i'^{-1} g_i(a) = 1$, let $\phi_1(\mathfrak{G}_R) = \mathfrak{G}_{\bar{R}}$, given by $\phi_1(a_i) = a_i$, be the corresponding isomorphism of $\mathfrak{G}_R$, and let $\bar{\mathbf{r}}^3 = T\{\phi_1(\mathbf{r}^3)\}$. Then $\bar{\mathbf{r}}^3 = \| \phi_1(\mathbf{r}^3), 0 \|$, where 0 indicates the columns which correspond to the new relations. The latter can obviously be transformed into the relations $a_i^{-1} f_i(a') = 1$ by transformations of the form (4.5), and these transformations leave $\bar{\mathbf{r}}^3$ unaltered, since the new columns consist entirely of zeros. We now transform $R_\lambda(a_1, \cdots, a_k) = 1$ into $R_\lambda'(a_1', \cdots, a_k') = 1$, for each $\lambda = 1, \cdots, \alpha_2$, by substituting $f_i(a')$ for a_i throughout the products R_λ. This operation is the resultant of transformations of the form

$$x a_i y' \to x a_i y' y'^{-1} a_i^{-1} f_i y' \qquad \text{or} \qquad x a_i^{-1} y' \to x f_i^{-1} a_i x^{-1} x a_i^{-1} y',$$

where $f_i = f_i(a')$, $x a_i^{\epsilon} y = R_\lambda$, for some $\lambda = 1, \cdots, \alpha_2$, and y' is the result of substituting $f_i(a')$ for a_i throughout the product y. Therefore it follows without difficulty that the transformation $R_\lambda \to R_\lambda'$ is the resultant of transformations of the form (4.5). Moreover the columns of $\mathbf{r}^3$, which correspond to the original relations $R_\lambda = 1$, are unaltered by these transformations.[35] Finally we discard

[35] As explained above, in the definition of $T_2(\mathbf{r}^3)$, it follows from the condition $\mathbf{r}^3\mathbf{r}^2 = 0$, and the form of the relations $a_i^{-1} f_i(a') = 1$, that the remaining (zero) columns also are ultimately left unaltered.

the generators $a_i, \cdots, a_k$ and the relations $a_i^{-1} f_i(a') = 1$ by transformations of the form T_2. Then $\phi_1(\mathfrak{r}^p)$ $(p = 4, \cdots, n)$ is transformed into $\phi_2(\bar{\mathfrak{r}}^p) = \phi_2\phi_1(\mathfrak{r}^p)$, where $\phi_2(a_i') = a_i'$, $\phi_2(a_i) = f_i(a')$, and $\phi_2\phi_1 = \phi$, since $\phi_2\phi_1(a_i) = \phi_2(a_i) = f_i(a')$. Similarly $\mathfrak{r}^3$ contracts into $\phi(\mathfrak{r}^3) = \mathfrak{r}'^3$ on discarding the (zero) columns, which correspond to the relations $a_i^{-1} f_i(a') = 1$. Therefore $(\mathfrak{r}, R)$ is L-equivalent to $(\mathfrak{r}', R')$ under the isomorphism ϕ.

What we have already proved, with $f_i(a') = a_i'$, allows us to 're-write' a_i as a_i' if $a_i' \neq a_j^{\pm 1}$, for any i, j and even if the sets a_i' and $a_i^{\pm 1}$ overlap. For we can first rewrite a_i as b_i and then re-write b_i as a_i', where $b_i^{\pm 1} \neq a_j$ or a_j'. Therefore, in completing the proof, we may assume that $a_i' = a_i$, in which case $a_i \to f_i(a)$ is an automorphism of $\mathfrak{F}_k$. Moreover we may re-write a_i' as a_i in what we have already proved. Therefore the theorem follows from the fact that the automorphisms of $\mathfrak{F}_k$ are generated by automorphisms of the form $a_i \to f_i(a)$, where[36] $f_p(a) = a_p^{-1}$ or $a_p a_q$, $f_i(a) = a_i$ $(q, i \neq p)$.

THEOREM 8. *If $\mathbf{K}_1^n$ and $\mathbf{K}_2^n$ have the same nucleus, then any natural systems for $\mathbf{K}_1^n$ and $\mathbf{K}_2^n$ are L-equivalent. If $\mathbf{K}_1^n$ and $\mathbf{K}_2^n$ are of the same homotopy type, then their natural systems are L^*-equivalent.*

This will obviously follow from the theorems 1 and 2, the arguments in Q.J., and S.S., theorem 13, when we have shown that two natural systems for the same complex $\mathbf{K}^n$, are L-equivalent. Let R and R' be two natural systems of generators and relations for $\pi_1(\mathbf{K}^n)$, and let $\phi(\mathfrak{G}_R) = \mathfrak{G}_{R'}$ be the isomorphism such that a given element in $\pi_1(\mathbf{K}^n)$ is represented by $x \in \mathfrak{G}_R$ and $\phi(x) \in \mathfrak{G}_{R'}$ After a change of basis, if necessary, we may obviously assume that $\mathfrak{r}'^p = \phi(\mathfrak{r}^p)$ $(p = 3, \cdots, n)$, where $\mathfrak{r}^n, \cdots, \mathfrak{r}^3$ and $\mathfrak{r}'^n, \cdots, \mathfrak{r}'^3$ are the matrices in natural systems, $(\mathfrak{r}, R)$ and $(\mathfrak{r}', R')$, for $\mathbf{K}^n$. Let $\mathbf{V}, \mathbf{V}' \subset \mathbf{K}^1$ be the trees in terms of which R and R' are defined. Then we may also assume that $\mathbf{V}' = \mathbf{V} - \mathbf{E}_1^1 + \mathbf{E}_2^1$, where $\mathbf{E}_1^1 \subset V$, $\mathbf{E}_2^1 \subset \mathbf{K}^1 - V$, and $\mathbf{E}_1^1$ separates the extremities of $\mathbf{E}_2^1$. For one can interchange any two trees, which are sub-complexes of $\mathbf{K}^1$ and contain $\mathbf{K}^0$, by reiterating transformations of the form $\mathbf{V} \to \mathbf{V}'$, where $\mathbf{V}$ and $\mathbf{V}'$ are thus related.[37]

On these assumptions let us, without altering the notation, shrink into a point each of the two trees into which $\mathbf{V}$ is separated by $\mathbf{E}_1^1$. Then $\mathbf{K}^1$ will consist of two vertices $\mathbf{E}_1^0$, $\mathbf{E}_2^0$, which are joined by certain 1-cells $\mathbf{E}_1^1, \mathbf{E}_2^1, \cdots, \mathbf{E}_p^1$, and of circuits $\mathbf{E}_{h1}^1, \cdots, \mathbf{E}_{hm_h}^1$ $(h = 1, 2)$ beginning and ending at $\mathbf{E}_h^0$. By theorem 7 we may orient $\mathbf{E}_i^1$ as we please, and we take $\mathbf{E}_1^0$ to be its first point. Let us denote $\mathbf{E}_i^1$ by s_i $(i = 1, \cdots, p)$ and $\mathbf{E}_{h\rho}^1$ by $\sigma_{h\rho}$. Then, as in (3.3), any segment on $\mathbf{K}'$ may be represented by a word

$$t_1^{\epsilon_1} \cdots t_q^{\epsilon_q}, \tag{4.6}$$

where $t_\lambda = s_{i_\lambda}$, or $\sigma_{h_\lambda \rho_\lambda}$. Let us now write $\sigma_{2\rho} = \sigma_\rho$. I say that, if (4.6) represents a circuit which begins and ends at $\mathbf{E}_1^0$, then, regarded as an element

[36] J. Nielsen, Math. Annalen, 79 (1919), 269–72.

[37] Cf. p. 105 of K. Reidemeister, Abh. Hamb. Sem., 11 (1935), 102–9.

in the free group, which is freely generated by s_i , σ_ρ , σ_{1a} , it is unaltered when we replace s_j by $s_j s_i^{-1}$ and σ_ρ by $s_i \sigma_\rho s_i^{-1}$, for each $j = 1, \cdots, p$, $\rho = 1, \cdots, m_2$, and any specified value of i. For each letter of the form s_j or $\sigma_\rho^{\pm 1}$ represents a segment which terminates at $\mathbf{E}_2^0$. Since the segment (4.6) terminates at $\mathbf{E}_1^0$ such a letter is followed by s_l^{-1} or by $\sigma_\tau^{\pm 1}$ for some value of l or ρ. Similarly s_j^{-1} or $\sigma_\rho^{\pm 1}$ is preceded by s_l or $\sigma_\tau^{\pm 1}$, since (4.6) begins at $\mathbf{E}_1^0$. Therefore, taking $\mathbf{E}_1^0$ as base point, we have $F(\mathbf{E}_\lambda^2) = R_\lambda(s_j , \sigma_\rho , \sigma_{1a})$, where

$$R_\lambda(s_j , \sigma_\rho , \sigma_{1a}) = R_\lambda(s_j s_i^{-1}, s_i \sigma_\rho s_i^{-1}, \sigma_{1a}),$$

in which the equality refers to the free group generated by $s_i , \sigma_\rho , \sigma_{1a}$. Taking $i = 1$ we have

$$F(\mathbf{E}_\lambda^2) = R_\lambda(1, a_j , b_\rho , \sigma_{1a}) \qquad (j = 2, \cdots, p), \tag{4.7}$$

where $a_j = s_j s_1^{-1}$, $b_\rho = s_1 \sigma_\rho s_1^{-1}$, and the system R, associated with $\mathbf{V}$, consists of the generators $a_2 , \cdots , a_p , b_\rho , \sigma_{1a}$, subject to the relations $R_\lambda = 1$. Similarly the system R', associated with $\mathbf{V}'$, is generated by $a_1' , a_3' , \cdots , a_p' , b_\rho' , \sigma_{1a}$, subject to the relations $R_\lambda' = 1$, where $a_t' = s_t s_2^{-1}$ $(t = 1, 3, \cdots, p)$, $b_\rho' = s_2 \sigma_\rho s_2^{-1}$ and

$$R_\lambda' = R_\lambda(a_1' , 1, a_\beta' , b_\rho' , \sigma_{1a}) \qquad (\beta = 3, \cdots, p). \tag{4.8}$$

On writing $s_2 = 1$ in the equation

$$R_\lambda(1, s_2 s_1^{-1}, s_\beta s_1^{-1}, s_1 \sigma_\rho s_1^{-1}, \sigma_{1a}) = R_\lambda(s_1 , s_2 , s_\beta , \sigma_\rho , \sigma_{1a}),$$

we have

$$R_\lambda(1, s_1^{-1}, s_\beta s_1^{-1}, s_1 \sigma_\rho s_1^{-1}, \sigma_{1a}) = R_\lambda(s_1 , 1, s_\beta , \sigma_\rho , \sigma_{1a}). \tag{4.9}$$

On comparing (4.7) with (4.8), it follows from (4.9) that R' is the image of R in the transformation given by

$$a_2 \to a_1'^{-1}, \qquad a_\beta \to a_\beta' a_1'^{-1}, \qquad b_\rho \to a_1' b_\rho' a_1'^{-1}, \qquad \sigma_{1a} \to \sigma_{1a} . \tag{4.10}$$

The product (4.6) is transformed into an element $x \in \mathfrak{G}_R$ if we write $s_j = a_j s_1$, $\sigma_\rho = s_1^{-1} b_\rho s_1$ and cancel all the factors $s_1^{\pm 1}$, and into $\phi_1(x) \in \mathfrak{G}_{R'}$ if we write $s_j = a_j' s_2$, $\sigma_\rho = s_2^{-1} b_\rho' s_2$ and cancel $s_2^{\pm 1}$, where $\phi_1(\mathfrak{G}_R) = \mathfrak{G}_{R'}$ is the isomorphism given by (4.10). Therefore ϕ_1 is the isomorphism ϕ, and since $\mathbf{r}'^p = \phi(\mathbf{r}^p)$ the theorem follows from theorem 7.

As in theorem 3 it is clear that the isomorphism $\psi\{\pi_1(\mathbf{K}_1^n)\} = \pi_1(\mathbf{K}_2^n)$, which is expressed algebraically by the isomorphism $\phi(\mathfrak{G}_R) = \mathfrak{G}_{R'}$, may be taken from the class which is determined by some homotopy isomorphism $f(\mathbf{K}_1^n) \subset \mathbf{K}_2^n$. The distinction between ϕ and ψ is analogous to one which is familiar in coordinate geometry, where an algebraic transformation may represent either a transformation of coordinates or a point transformation, according to the context.

We now prove an analogous theorem to theorem 8, in which the parts played by the algebraic and the geometric equivalence are interchanged. In general the analogy is not complete. For two complexes with the same natural system

need not even be of the same homotopy type.[38] Moreover, there is no reason to suppose that every system $(\mathfrak{r}, R)$ can be realized geometrically as a natural system of a complex. The theorem is

THEOREM 9. *Any system* $(\mathfrak{r}', R')$, *which is* X*-equivalent* ($X = L$ *or* L^*) *to a natural system,* $(\mathfrak{r}, R)$, *for a given complex* $\mathbf{K}^n$, *is a natural system for some complex of the same homotopy type as* $\mathbf{K}^n$, *and for some complex with the same nucleus of* $X = L$.

First let $X = L$. Then $(\mathfrak{r}, R)$ can be transformed into $(\mathfrak{r}', R')$ by a sequence of elementary expansions and contractions, according to lemma 4. Therefore the theorem will follow inductively, for $X = L$, when we have proved it in case $(\mathfrak{r}', R')$ is obtained from $(\mathfrak{r}, R)$ by a single elementary expansion or contraction.

Let $(\mathfrak{r}', R')$ be obtained from $(\mathfrak{r}, R)$ by an elementary expansion. First let $(\mathfrak{r}', R') = T_1(\mathfrak{r}, R)$, and let a_0 be the new generator and $xa_0^\epsilon y = 1$ the new relation. Let $s, t \subset \mathbf{K}^1$ be oriented circuits, beginning and ending at the base point $\mathbf{E}^0 \epsilon \mathbf{K}^0$, which are represented by x and y respectively. Let $\mathbf{E}_0^1$ be a new 1-cell, beginning and ending at $\mathbf{E}^0$ which shall correspond to the new generator a_0. Let $\mathbf{E}_0^2$ be a new 2-cell such that $F(\mathbf{E}_0^2) \subset \mathbf{K}^1 + \mathbf{E}_0^1$ is given by $xa_0^\epsilon y$, with a suitable base point. Then $(\mathfrak{r}', R')$ is a natural system for $\mathbf{K}_1^n = \mathbf{K}^n + \mathbf{E}_0^2$, and assuming that $F(\mathbf{E}_0^2)$ is simple in $\mathbf{E}_0^1$, as we obviously may, $\mathbf{K}^n$ expands into $\mathbf{K}_1^n$. Secondly let $(\mathfrak{r}', R') = T_3(\mathfrak{r}, R)$ and let $R_0 = 1$ be the new relation, where R_0 is given by (4.1). With the same notation as in §3, there is obviously a singular 2-cell $\tilde{\mathbf{G}}^2 \subset \tilde{\mathbf{K}}^2$, whose boundary, $F(\tilde{\mathbf{G}}^2) \subset \tilde{\mathbf{K}}^1$, is represented by the product R_0, such that the relative 2-cycle, mod. $\tilde{\mathbf{K}}^1$, determined by $\tilde{\mathbf{G}}^2$ is

$$\epsilon_1 x_1 \tilde{\mathbf{E}}_{\lambda_1}^2 + \cdots + \epsilon_q x_q \tilde{\mathbf{E}}_{\lambda_q}^2 = c_1 \tilde{\mathbf{E}}_1^2 + \cdots + c_{\alpha_2} \tilde{\mathbf{E}}_{\alpha_2}^2 \qquad (c_i \epsilon \mathfrak{R}).$$

Let $\mathbf{E}_0^2$ be a new membrane whose boundary is given by $F(\mathbf{E}_0^2) = u\{F(\tilde{\mathbf{G}}^2)\} \subset \mathbf{K}^1$, let $\mathbf{G}^2 = u(\tilde{\mathbf{G}}^2) \subset \mathbf{K}^2$ and let $\mathbf{E}_0^3$ be a new membrane bounded by the spherical map $\epsilon(\mathbf{E}_0^2 - \mathbf{G}^2)$ ($\epsilon = \pm 1$). Then $F(\mathbf{E}_0^3)$ is simple in $\mathbf{E}_0^2$, whence $\mathbf{K}^n$ expands into $\mathbf{K}_1^n = \mathbf{K}^n + \mathbf{E}_0^3$. Assuming that $\tilde{\mathbf{K}}^n$ is imbedded in $\tilde{\mathbf{K}}_1^n$, let $\tilde{\mathbf{E}}_0^2 \epsilon u^{-1}(\mathbf{E}_0^2) \subset \tilde{\mathbf{K}}_1^2$, and $\tilde{\mathbf{E}}_0^3 \epsilon u^{-1}(\mathbf{E}_0^3) \subset \tilde{\mathbf{K}}_1^3$ be chosen so that $F(\tilde{\mathbf{E}}_0^2) = F(\tilde{\mathbf{G}}^2)$ and $F(\tilde{\mathbf{E}}_0^3) = \epsilon x(\tilde{\mathbf{E}}_0^2 - \tilde{\mathbf{G}}^2)$, where

$$\mathfrak{r}'^3 = \begin{Vmatrix} \epsilon x & -\epsilon x \mathfrak{c} \\ 0 & \mathfrak{r}^3 \end{Vmatrix},$$

as in (4.4). Then

$$\partial \tilde{\mathbf{E}}_0^3 = \epsilon x \tilde{\mathbf{E}}_0^2 - \epsilon x(c_1 \tilde{\mathbf{E}}_1^2 + \cdots + c_{\alpha_2} \tilde{\mathbf{E}}_{\alpha_2}^2),$$

whence $(\mathfrak{r}', R')$ is a natural system for $\mathbf{K}_1^n$. An elementary expansion of the form (2.5) with $p > 3$, can be copied in the same way.

Let $(\mathfrak{r}', R') = T_2(\mathfrak{r}, R)$, and let a_0 be the generator and $xa_0^\epsilon y = 1$ the relation to be removed. Since $a_0^{\pm 1}$ does not appear in any other relation we may assume,

[38] Cf. note III_3.

after a trivial deformation of $\mathbf{K}^2$ if necessary, that $F(\mathbf{E}_i^2) \subset \mathbf{K}_1^1$ $(i = 1, \cdots, \alpha_2)$, where $\mathbf{E}_0^1, \cdots, \mathbf{E}_{\alpha_1}^1$ are the 1-cells of $\mathbf{K}^n$, and

$$\mathbf{K}_1^1 = \mathbf{K}^0 + \mathbf{E}_1^1 + \cdots + \mathbf{E}_{\alpha_1}^1 .$$

We also assume that $F(\mathbf{E}_0^2)$ is simple in $\mathbf{E}_0^1$, where $\mathbf{E}_0^1$ and $\mathbf{E}_0^2$ are the cells corresponding to a_0 and $xa_0^\epsilon y = 1$. Then $\mathbf{K}^2$ contracts into $\mathbf{K}_1^2$, where

$$\mathbf{K}_1^2 = \mathbf{K}_1^1 + \mathbf{E}_1^2 + \cdots + \mathbf{E}_{\alpha_2}^2 .$$

It follows from an obvious induction that $\mathbf{K}^n$ may be transformed by a series of trivial deformations into a complex of the form $\mathbf{K}_1^n + \mathbf{E}_0^2$, where neither $\mathbf{E}_0^2$ nor $\mathbf{E}_0^1$ has an inner point in common with $\mathbf{K}_1^n$. In other words we push the remaining cells in $\mathbf{K}^n$ off the interiors of $\mathbf{E}_0^1$ and $\mathbf{E}_0^2$. Then $\mathbf{K}_1^n + \mathbf{E}_0^2$ contracts into $\mathbf{K}_1^n$, and $(\mathfrak{r}', R')$ is a natural system for $\mathbf{K}_1^n$.

Let $(\mathfrak{r}, R) \to (\mathfrak{r}', R')$ be due to a contraction of $\mathfrak{r}^p$, for some $p > 3$, in which the 0^{th} row and column are removed. Then $\tilde{\mathbf{E}}_0^{p-1}$ appears in $\partial\tilde{\mathbf{E}}_0^p$ with the coefficient $\pm x$, for some $x \in \mathfrak{G}_R$, and with the coefficient zero in $\partial\tilde{\mathbf{E}}_i^p$ $(i = 1, \cdots, \alpha_p)$. Let $\mathbf{A}^{p-1} \in \mathbf{E}_0^{p-1}$ be an open simplex, whose closure is internal to $\mathbf{E}_0^{p-1}$, and let $y\tilde{\mathbf{A}}^{p-1}$ be its image in $y\tilde{\mathbf{E}}_0^{p-1}$, for each $y \in \mathfrak{G}_R$. Then it follows from a standard theorem [39] that $F(\tilde{\mathbf{E}}_i^p)$ $(i = 1, \cdots, \alpha_p)$ is homotopic in $\tilde{\mathbf{K}}^{p-1}$ to a map in which $y\tilde{\mathbf{A}}^{p-1}$ is uncovered for each $y \in \mathfrak{G}_R$. Also $F(\tilde{\mathbf{E}}_0^p)$ is homotopic to a map which is simple in $x\tilde{\mathbf{A}}^{p-1}$ and in which $y\tilde{\mathbf{A}}^{p-1}$ is uncovered if $y \neq x$. It follows that $F(\mathbf{E}_0^p)$ is homotopic in $\mathbf{K}^{p-1}$ to a map which is simple in $\mathbf{E}_0^{p-1}$, and that $F(\mathbf{E}_i^p)$ $(i = 1, \cdots, \alpha_p)$ is homotopic to a map in which $\mathbf{A}^{p-1}$ is uncovered, and hence to a map in

$$\mathbf{K}_1^{p-1} = \mathbf{K}^{p-2} + \mathbf{E}_1^{p-1} + \cdots + \mathbf{E}_{\alpha_{p-1}}^{p-1} .$$

The same is true if $p = 3$ and $(\mathfrak{r}', R') = T_4(\mathfrak{r}, R)$. For since the relation $R_0 = 1$, which is to be discarded, is redundant, it follows that $\tilde{\mathbf{K}}^2 - u^{-1}(\mathbf{A}^2)$ is simply connected, where $\mathbf{A}^2 \in \mathbf{E}_0^2$. Therefore the above argument, which fails, in general, unless $p - 1 > 2$, is valid here if $p - 1 = 2$. In either case it now follows from the argument used where $(\mathfrak{r}', R') = T_2(\mathfrak{r}, R)$, that $(\mathfrak{r}', R')$ is a natural system for a complex $\mathbf{K}_1^n$, which is obtained from $\mathbf{K}^n$ by a series of trivial deformations, followed by a contraction. Therefore the theorem is established in case $X = L$.

In proving theorem 17 in S.S. it was shown that two n-dimensional complexes are of the same homotopy type if they have the same $(n + 1)$-group. Therefore the theorem will follow, in case $X = L^*$, when we have shown that (x', R') is a natural system for an (n-dimensional) complex with the same $(n + 1)$-group as $\mathbf{K}^n$. Let empty matrices $\mathfrak{r}^p$, $\mathfrak{r}'^p$ $(p = n + 1, n + 2)$, with α_p, α_p', rows and α_{p-1}, α_{p-1}' columns $(\alpha_{n+2} = \alpha_{n+2}' \geqq 0, \alpha_{n+1} = \alpha_{n+1}' = 0, \alpha_n, \alpha_n' =$ number of

[39] Cf. P. Alexandroff and H. Hopf, (loc. cit.), p. 503, or pp. 99–100 or S. Lefschetz, Fund. Math., 27 (1936), 94–115.

rows in $\mathbf{r}^n$, $\mathbf{r}'^n$), be included in the systems $(\mathbf{r}, R)$ and $(\mathbf{r}', R')$. The augmented system $(\mathbf{r}, R)$ is a natural system for any complex of the form

$$\mathbf{K}^{n+2} = \mathbf{K}^n = \mathbf{E}_1^{n+2} + \cdots + \mathbf{E}_{\alpha_{n+2}}^{n+2},$$

where $F(\mathbf{E}_\lambda^{n+2}) \subset \mathbf{K}^n$. Since the original systems $(\mathbf{r}, R)$ and $(\mathbf{r}', R')$ are L^*-equivalent, it follows from the addendum to theorem 2 that the augmented systems are L-equivalent, provided α_{n+2} is suitably chosen, which we assume to be the case. Therefore, by what we have already proved, the augmented system $(\mathbf{r}', R')$ is a natural system for a complex $\mathbf{K}_1^{n+2}$, which has the same nucleus, and hence the same $(n+1)$-group as $\mathbf{K}^{n+2}$. It follows from the definition in §3 above that $\mathbf{K}^{n+2}$, and hence $\mathbf{K}_1^{n+2}$ has the same $(n+1)$-group as $\mathbf{K}^n$. In general, if $(\mathbf{r}, R)$ is a natural system for a complex $\mathbf{K}^n$, then $(\mathbf{r}^p, \cdots, \mathbf{r}^3, R)$ $(3 \leqq p \leqq n)$ is a natural system for $\mathbf{K}^p$. Therefore the original system $(\mathbf{r}', R')$ is a natural system for $\mathbf{K}_1^n$. Since $\alpha'_{n+1} = 0$ the complex $\mathbf{K}_1^n$ is obtained from $\mathbf{K}_1^{n+2}$ by removing the interiors of the $(n+2)$-cells, and hence has the same $(n+1)$-group as $\mathbf{K}_1^{n+2}$. Therefore $\mathbf{K}_1^n$ has the same $(n+1)$-group as $\mathbf{K}^n$, and the proof is complete.

If $X = L$ the order[40] of the deformation $D(\mathbf{K}^n) = \mathbf{K}_1^n$, in theorem 9, need not exceed $n+1$. Also the formal deformations involved in theorem 5 are of order $n+1$ at most. Therefore we have the corollary to theorems 5, 8 and 9:

COROLLARY. *If two complexes $\mathbf{K}_0^n$ and $\mathbf{K}^n$ have the same nucleus and either $n = 3$, or $n > 3$ and $\pi_r(\mathbf{K}^n) = 0$ for $r = 2, \cdots, n-2$, then $\mathbf{K}^n = D(\mathbf{K}_0^n)$, where the order of D does not exceed $n+1$.*

For let $(\mathbf{r}, R)$ and $(\mathbf{r}', R')$ be natural systems for $\mathbf{K}_0^n$ and $\mathbf{K}^n$ respectively. By theorem 8, $(\mathbf{r}', R')$ and $(\mathbf{r}, R)$ are L-equivalent. It follows from theorem 9 that $(\mathbf{r}', R')$ is a natural system for a complex $\mathbf{K}_1^n = D_1(\mathbf{K}_0^n)$, and from theorem 5, that $\mathbf{K}^n = D_2(\mathbf{K}_1^n)$, where D_1 and D_2 are of order $n+1$ at most.

In $\pi_1(\mathbf{K}^n)$ satisfies the condition explained in S.S., §11, it follows from S.S., theorem 21, that the corollary remains valid with "have the same nucleus" replaced by "are of the same homotopy type." For example, an n-dimensional simplicial complex, which is an absolute retract, expands into a collapsible complex of at most $n+1$ dimensions.

5. In this section we prove the theorem on lens spaces, which was announced in §1. We first prove a lemma concerning the ring of a cyclic group. Let $\mathfrak{G}$ be a cyclic group of order m, let $\mathfrak{R}$ be its group ring, and let $\sigma = 1 + \cdots + x^{m-1}$, where x generates $\mathfrak{G}$. We shall describe an element $\alpha \in \mathfrak{R}$ as *regular* mod. σ if, and only if, there is an element $\beta \in \mathfrak{R}$, such that

$$\alpha\beta = 1 + k\sigma, \tag{5.1}$$

for some value of k. Since $x^p\sigma = \sigma$ it follows that $\lambda\sigma = S(\lambda)\sigma$, for any $\lambda \in \mathfrak{R}$,

[40] By the order of a formal deformation is meant the maximum dimensionality of the cells introduced or removed.

where $S(\lambda)$ denotes the algebraic sum of the coefficients in λ. Therefore we assume that k, in (5.1), is an integer. The lemma is:

LEMMA 6. *If α is regular* mod. σ, *and if $S(\alpha) = \pm 1$, then α is regular.*

Writing $S(\alpha) = \epsilon$ we have $\alpha(\beta - \epsilon k\sigma) = 1 + k\sigma - k\sigma = 1$.

We also recall that, if $\lambda(x^q - 1) = 0$, where q is prime to m, then $\lambda \equiv 0$ (mod. σ), meaning that $\lambda = k\sigma$, for some value of k. If $\lambda(x^q - 1)^n = 0$, where $n > 1$, we have $\lambda(x^q - 1)^{n-1} = k\sigma$, whence $km\sigma = k\sigma^2 = \lambda(x^q - 1)^{n-1}\sigma = 0$. It follows from induction on n that, if $\lambda(x^q - 1)^n = 0$, for any $n > 0$, then $\lambda \equiv 0$ (mod. σ).

THEOREM 10. *Two lens spaces of types (m, q) and (m, r) are of the same homotopy type if, and only if,*[41] $r \equiv \pm l^2 q$ (mod. m), *for some value of l.*

Let M_1^3 be a lens space of type (m, q), which we assume to be represented by a lens model of this type. Then $(\mathbf{r}, R)$ is a natural system for M_1^3, where $\mathbf{r}$ consists of the matrix $\mathbf{r}^3 = \| x^q - 1 \|$, and R consists of a single generator x, and the relation $x^m = 1$. The corresponding incidence matrices are

$$\| x^q - 1 \|, \qquad \| \sigma \|, \qquad \| x - 1 \|.$$

If M_1^3 is of the same homotopy type as a lens space of type (m, r) it follows from theorem 8 and 6, and lemma 5, that

$$(x^{ql} - 1)(x^l - 1) = \alpha(x^r - 1)(x - 1), \tag{5.2}$$

for some l which is prime to m, and some regular element $\alpha \in \mathfrak{R}$. Let $\sigma_n = 1 + \cdots + x^{n-1}$ $(1 \leqq n \leqq m)$. Then $x^n - 1 = (x - 1)\sigma_n$, and it follows from (5.2) and a previous observation that

$$\sigma_{ql}\sigma_l \equiv \alpha\sigma_r \qquad \text{(mod. } \sigma\text{)}.$$

On writing $x = 1$ we have $ql^2 \equiv \pm r$ (mod. m), since $S(\alpha) = \pm 1$, the element α being regular.

Conversely, let

$$r \equiv \pm l^2 q \qquad \text{(mod. } m\text{)}. \tag{5.3}$$

If a and b are given integers and if $| a, m | = 1$, where $| c, d |$ denotes the positive H.C.F. of c and d, there is an element $\xi \in \mathfrak{R}$ such that[42] $\xi(x^a - 1) = x^b - 1$. If also $| b, m | = 1$ there is an η such that $\eta(x^b - 1) = x^a - 1$. Then

$$\begin{aligned} x^b - 1 &= \xi(x^a - 1) \\ &= \xi\eta(x^b - 1), \end{aligned}$$

whence $\xi\eta \equiv 1$ (mod. σ). Therefore $| b, m | = 1$ implies that ξ is regular mod. σ. Since $| r, m | = 1$ there is a $\xi \in \mathfrak{R}$, such that

$$\xi(x^r - 1) = x^{ql} - 1. \tag{5.4}$$

[41] Of course the necessity of this condition follows from the standard theorems on the Kronecker index and the fact that, if two manifolds are of the same homotopy type, then each is the image of the other in a map of degree 1.

[42] See p. 107 of Reidemeister, Abh. Hamb. Sem., 11 (1935), 102-9.

Also it follows from (5.3) that $| ql, m | = | l, m | = 1$. Therefore ξ is regular mod. σ, and so is $\sigma_l(x-1) = x^l - 1$. Therefore $\beta = \xi\sigma_l$ is regular mod. σ. It follows from (5.4) that

$$\xi\sigma_r \equiv \sigma_{ql} \qquad \text{(mod. } \sigma\text{)},$$

whence $rS(\xi) \equiv ql$ (mod. m). Since $| ql, m | = 1$ it follows from (5.3) that $lS(\xi) \equiv \pm 1$ (mod. m), whence $S(\beta) = lS(\xi) = \pm 1 + tm$, for some integral value of t. Therefore $S(\alpha) = \pm 1$, where $\alpha = \beta - t\sigma$. Since β is regular mod. σ, so is α, and it follows from lemma 6 that α is regular. Since $(\beta - \alpha)(x^r - 1) = t\sigma(x^r - 1) = 0$, it follows from (5.4) that

$$\begin{aligned} \alpha(x^r - 1) &= \beta(x^r - 1) \\ &= \xi(x^r - 1)\sigma_l \\ &= (x^{ql} - 1)\sigma_l . \end{aligned} \tag{5.5}$$

Since α is regular the natural system $(\| x^r - 1 \|, R)$, for a lens space M^3_2 of type (m, r), is L^*-equivalent to the system $\{\| \alpha(x^r - 1) \|, R\}$. By theorem 9 there is therefore a membrane complex $\mathbf{K}^3_2$, of the same homotopy type as M^3_2, for which $\{\| \alpha(x^r - 1) \|, R\}$ is a natural system. Therefore the theorem will follow from (5.5) and theorem 5 when we have shown that there is a membrane complex $\mathbf{K}^3_1$, of the same homotopy type as M^3_1, for which $\{\| (x^{ql} - 1)\sigma_l \|, R\}$ is a natural system. It will therefore follow from theorem 9 when we have shown that the system $(\| x^q - 1 \|, R)$ is L^*-equivalent to $\{\| (x^{ql} - 1)\sigma_l \|, R\}$ under some automorphism of $\mathfrak{G}_R$. Actually we shall show that the two systems are L-equivalent under the automorphism in which x corresponds to x^l.

Starting with $(\| x^q - 1 \|, R)$ we extend R to a system R' by the addition of a new generator y and the new relation $yx^{-l'} = 1$, where $ll' \equiv 1$ (mod. m). Such an l' exists since $| l, m | = 1$. At the same time we transform $\| x^q - 1 \|$ into the matrix $\| x^q - 1, 0 \|$. Let $x_1 = yx^l$, $y_1 = x$. Since $x = y_1$, $y = x_1 y_1^{-l}$, it follows that x_1, y_1 also generate the free group $\mathfrak{F}_2$, which is freely generated x, y, whence the transformation $x \to yx^l$, $y \to x$ determines an automorphism of $\mathfrak{F}_2$. By theorem 7, the system $(\| x^q - 1, 0 \|, R')$ is L-equivalent to $\{\| (yx^l)^q - 1, 0 \|, R''\}$ under the isomorphism given by $x \to yx^l$, $y \to x$, where R'' is generated by x, y, subject to the relations

$$(yx^l)^m = 1, \qquad x(yx^l)^{-l'} = 1. \tag{5.6}$$

I say that the relations (5.6) can be reduced to $x^m = 1$ by transformations of the form (4.5), followed by one of the form T_2. For, since a relation of the form $x^a x_1^b$ may be replaced by $x_1^b x^a$, we may, for the purposes of the reduction, allow x^a to commute with x_1^b. Let us write

$$A = mx_1, \qquad B = x - l'x_1,$$

where $x_1 = yx^l$ and the addition is commutative. Then

$$l'A + mB = mx, \qquad kA + lB = lx - x_1,$$

where $km = ll' - 1$. Since $ll' - km = 1$ the transformation given by

$$A' = l'A + mB, \qquad B' = kA + lB$$

is the resultant of a sequence of elementary transformations of the form $(A, B) \to (A \pm B, B)$ or $(A, B \pm A)$. Therefore a similar sequence of the transformations (4.5), punctuated with suitable transformations of the form $x^a x_1^b \to x_1^b x^a$, will transform the relations (5.6) into the relations $x^m = 1$, $x^l x_1^{-1} = y^{-1} = 1$. On discarding the generator y and the relation $y^{-1} = 1$, it follows that the system $(\| x^q - 1 \|, R)$ is L-equivalent to $\{\| x^{ql} - 1)\lambda \|, R\}$, under the automorphism given by $x \to x^l$, where λ is the factor due to the transformation of the relations (5.6) into $x^m = 1$, $y^{-1} = 1$. It follows from theorem 6 that the sets of matrices

$$\| x^{ql} - 1 \|, \qquad \| \sigma \|, \qquad \| x^l - 1 \|$$

and

$$\| (x^{ql} - 1)\lambda \|, \qquad \| \sigma \|, \qquad \| x - 1 \|$$

are L-equivalent, and from lemma 5 that

$$\begin{aligned}(x^{ql} - 1)(x - 1)\lambda &= \pm x^p (x^{ql} - 1)(x^l - 1) \\ &= \pm x^p (x^{ql} - 1)(x - 1)\sigma_l ,\end{aligned}$$

for some value of p. Therefore

$$(x^{ql} - 1)\lambda = \pm x^p (x^{ql} - 1)\sigma_l .$$

On multiplying $(x^{ql} - 1)\lambda$ by $\pm x^{-p}$ it follows that $(\| x^q - 1 \|, R)$ is L-equivalent under the automorphism $x \to x^l$ to $\| (x^{ql} - 1)\sigma_l , R \|$, and the proof is complete.

6. We conclude with some notes. The sections to which they primarily refer are indicated by Roman numerals.

II_1. Let $\mathfrak{G}_0$ be the free product of an infinite cyclic group, generated by x, and a cyclic group of order two, generated by y. Let $\mathfrak{R}$ be the group ring of $\mathfrak{G}_0$, and let the group $\mathfrak{G}$, in terms of which elementary matrices are defined, consist of the elements $\pm x$, where $x \in \mathfrak{G}_0$. Finally let $\lambda = x(1 + y)$, $\mu = 1 - y$, where 1 stands for the unit element in $\mathfrak{R}$. Then $\lambda\mu = 0$, $\mu\lambda \neq 0$, and it may be verified that

$$\left\| \begin{matrix} 1 - \mu\lambda & 0 \\ 0 & 1 \end{matrix} \right\| = \mathbf{a}\mathbf{b}\mathbf{a}^{-1}\mathbf{b}^{-1},$$

where

$$\mathbf{a}^\epsilon = \left\| \begin{matrix} 1 & 0 \\ \epsilon\lambda & 1 \end{matrix} \right\|, \qquad \mathbf{b}^\epsilon = \left\| \begin{matrix} 1 & \epsilon\mu \\ 0 & 1 \end{matrix} \right\| \quad (\epsilon = \pm 1).$$

Therefore an expansion of a matrix may be elementary, although the original matrix is not.

II_2. *Concerning the condition* (2.3). First notice that this is a consequence of the, apparently, less restrictive condition in which $\mathbf{a}$ and $\mathbf{b}$ are to be elementary matrices. For if the matrices $\left\|\begin{matrix} 0 & \mathbf{e}_k \\ 0 & 0 \end{matrix}\right\|$ and $\left\|\begin{matrix} 0 & \mathbf{e}_l \\ 0 & 0 \end{matrix}\right\|$ have p rows and q columns, we can border them with p final rows, and q initial columns of zeros, at the same time replacing given regular matrices, $\mathbf{a}$ and $\mathbf{b}$, by the elementary matrices $\left\|\begin{matrix} \mathbf{a} & 0 \\ 0 & \mathbf{a}^{-1} \end{matrix}\right\|$ and $\left\|\begin{matrix} \mathbf{b}^{-1} & 0 \\ 0 & \mathbf{b} \end{matrix}\right\|$.

This condition is also satisfied if $\mathfrak{R}$ is a ring in which $\mathbf{ab} = \mathbf{e}_r$ implies that $\mathbf{a}$ is regular, where $\mathbf{a}$ is a square matrix of degree r. The following example shows that it is not always satisfied. Let $\mathfrak{R}_0$ be the free ring, freely generated by a, b, p, q, whose elements are finite sums of words in a, b, p, q. Let $\mathfrak{R}$ be the ring which is obtained from $\mathfrak{R}_0$ by adding a unit element e, and the relations

$$ex = xe = x \qquad (x = a, v, p, q \text{ or } e)$$

$$ap + bq = pa = qb = e, \qquad pb = qa = 0.$$

The ring $\mathfrak{R}$ does not reduce to the zero element alone. For any formal polynomial in e, a, b, p, q can be reduced by means of these relations, together with the simplifications $mx + nx = (m + n)x$ $(m, n = 0, \pm 1, \pm 2, \cdots)$, $x + 0y = x$ and $xzy + xz'y = x(z + z')y$, to a unique normal form, in which the letters e, a, b, p, q occur a minimum number of times.[43] This can be proved by enumerating a set of "elementary reductions," which, with their inverses, generate the transformations by which any two expressions for a given element of $\mathfrak{R}$ can be interchanged. One can then show that the condition for "random reduction" is satisfied. That is to say, if two reductions r_1 and r_2 are applicable to a polynomial P, then $t_1r_1(P) = t_2r_2(P)$, where t_i either reduces $r_i(P)$ or leaves it unaltered. The expression e cannot be reduced, whence $e \neq 0$ in $\mathfrak{R}$.

Let

$$\mathbf{c} = \left\|\begin{matrix} p & 0 & 0 \\ q & 0 & 0 \\ 0 & a & b \end{matrix}\right\|, \qquad \mathbf{c}' = \left\|\begin{matrix} a & b & 0 \\ 0 & 0 & p \\ 0 & 0 & q \end{matrix}\right\|.$$

Then it follows by direct calculation that $\mathbf{cc}' = \mathbf{c}'\mathbf{c} = \mathbf{e}_3$ and that

$$\mathbf{c}\left\|\begin{matrix} 0 & 0 & e \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix}\right\| = \left\|\begin{matrix} 0 & 0 & p \\ 0 & 0 & q \\ 0 & 0 & 0 \end{matrix}\right\| = \left\|\begin{matrix} 0 & e & 0 \\ 0 & 0 & e \\ 0 & 0 & 0 \end{matrix}\right\|\mathbf{c}'.$$

[43] E.g. a is to be counted once, not six times, in $6a$.

Notice that $\mathbf{c}$ is even an elementary $\mathfrak{G}$-matrix, where $\mathfrak{G}$ consists of the single element e. For it is transformed into the elementary matrix

$$\begin{Vmatrix} 0 & -e & 0 \\ 0 & 0 & -e \\ e & 0 & 0 \end{Vmatrix},$$

as follows

$$\rho_3 \to \rho_3 + a\rho_1 + b\rho_2 ; \qquad \kappa_2 , \kappa_3 \to \kappa_2 - \kappa_1 a, \kappa_3 - \kappa_1 b;$$

$$\rho_1 , \rho_2 \to \rho_1 - p\rho_3 , \rho_2 - q\rho_3 ,$$

where ρ_i and κ_i denote the i^{th} row and the i^{th} column of a given matrix.

III_1. Two spaces X and Y are of the same homotopy type if, and only if, there are maps $f(X) \subset Y$ and $g(Y) \subset X$ such that $gf(X) \subset X$ and $fg(Y) \subset Y$ are each homotopic to the identity. We shall describe f as a *homotopy isomorphism*. Assuming that X and Y are arcwise connected, a homotopy isomorphism $f(X) \subset Y$ determines a unique class of isomorphisms $\psi\{\pi_1(X)\} = \pi_1(Y)$, any one of which is given by $\psi = \psi_0 h$, where ψ_0 is a particular isomorphism in the class, and h is an inner automorphism of $\pi_1(X)$.

III_2. The theory of nuclei and m-groups can also be developed as follows. Let $\mathbf{E}^n$ now stand for a non-singular cell, bounded by a given map $f(S^{n-1}) \subset X$, where X is any topological space. We recall the definition of $X + \mathbf{E}^n$, given in a paper entitled "On adding relations to homotopy groups" which has been submitted to the Annals of Mathematics, and which will be referred to as R.H. Let $S^{n-1} = \dot{E}^n$, where E^n is an n-element which does not meet X, and let $\mathbf{E}^n$ be the interior of E^n. Then $X + \mathbf{E}^n$ consists of X and $\mathbf{E}^n$, each with their own topology, fastened together by the condition that an infinite sequence of points $p_1 , p_2 , \cdots \subset \mathbf{E}^n \subset X + \mathbf{E}^n$ shall converge to $q \in f(\dot{E}^n)$ if, and only if, the limit points of $p_1 , p_2 , \cdots \subset \mathbf{E}^n$ are all contained in $f^{-1}(q) \subset \dot{E}^n$. In other words we identify each point $p \in \dot{E}^n$ with $f(p) \in X$, without identifying any two points in $\mathbf{E}^n$. A map $f(S^n) \subset X + \mathbf{E}^n$ may be described as *simple* in $\mathbf{E}^n$ if, and only if, it is a homeomorphism throughout $f^{-1}(U)$, where U is some relatively open set in $\mathbf{E}^n$. Then a membrane complex $\mathbf{K}^n$, and the nucleus and m-group of $\mathbf{K}^n$, may be defined as in §3 above. Many of the theorems in S.S., notably theorems 11, 14 and 17, remain valid when reinterpreted in terms of these definitions. Theorem 11, for example, may be proved as in §3 above. In theorem 17 one may take the analogue of a simplicial map to be a map $f(\mathbf{K}_0) \subset \mathbf{K}_1$, such that $f(\mathbf{K}_0^p) \subset \mathbf{K}_1^p$, for each $p = 0, 1, \cdots$.

Let $\mathbf{E}^n$ be a non-singular cell, which is bounded by a simplicial map $f(S^{n-1})$ in a simplicial complex K. Then it is to be expected that $\bar{\mathbf{E}}^n$, the closure of $\mathbf{E}^n$ in $K + \mathbf{E}^n$, can be triangulated.[44] Let us assume that $\bar{\mathbf{E}}^n$ can be triangulated to form a simplicial complex C^n, such that $C^n - \mathbf{A}^n$ contracts into the sub-

[44] It seems probable that $\bar{\mathbf{E}}^n$ is homeomorphic to a simple membrane, as defined in S.S.

complex covering $F(\mathbf{E}^n) = f(S^{n-1})$, where $\mathbf{A}^n$ is some open simplex of C^n, whose closure is in the (open) cell $\mathbf{E}^n$. Then it is easily shown that any membrane complex can be transformed into a polyhedron by a series of trivial deformations, as defined in §3 above; also that two simplicial complexes have the same nucleus, or m-group, in the sense of S.S., if they have the same nucleus, or m-group, when treated as membrane complexes. If the above assumption is correct it follows that the introduction of membrane complexes does not involve anything essentially new. But, in some problems, the freedom from the combinatorial limitations of S.S. is likely to be an advantage.

III_3. In this note we indicate how the incidence matrices for a membrane complex can be defined in terms of R.H. Let $\mathbf{E}_i^q$ $(q = 0, \cdots, n; i = 1, \cdots, m_q)$ be the cells of a membrane complex $\mathbf{K}^n$, which may be defined either as in §3 or as in III_2 above. Let $F(\mathbf{E}_i^p)$ be joined to $\mathbf{E}_1^0$ by a segment in $\mathbf{K}^{p-1}$ $(p > 2)$, so as to determine an element $\alpha_i^{p-1} \in \pi_{p-1}(\mathbf{K}^{p-1})$. Then $\mathbf{r}^p$ is the matrix of the coefficients in

$$\psi(\alpha_\lambda^{p-1}) = \sum_{i=1}^{m_{p-1}} r_{\lambda i}^p e_i^{p-1}, \tag{6.1}$$

where e_i^{p-1} is the basis element of the modulus M, in R.H., which corresponds to $\mathbf{E}_i^{p-1}$. Theorem 5 above, and lemma 3 in R.H., enable us to describe the indeterminacy in the description of $\mathbf{K}^n$ by means of a natural system $(\mathbf{r}, R)$. The nucleus of $\mathbf{K}^3$ is uniquely determined by R and $\mathbf{r}^3$. When $\mathbf{K}^{p-1}$ is given, for $p > 3$, the elements $\alpha_1^{p-1}, \cdots, \alpha_{m_p}^{p-1}$ determine $\mathbf{K}^p$ up to a trivial deformation rel. $\mathbf{K}^{p-1}$, and $\mathbf{r}^p$ is given by (6.1). In passing from the description of $\mathbf{E}_1^p, \cdots, \mathbf{E}_{m_p}^p$ by means of $\alpha_1^{p-1}, \cdots, \alpha_{m_p}^{p-1}$, to their description by means of $\mathbf{r}^p$, we throw away the sub-group $\pi_{p-1}^0 \subset \pi_{p-1}(\mathbf{K}^{p-1})$, whose elements have representative maps $f(S^{p-1}) \subset \mathbf{K}^{p-2}$.

For example, let $\mathbf{K}_i^4 = \mathbf{E}_i^2 + \mathbf{E}_i^4$ $(i = 0, 1)$, where $F(\mathbf{E}_i^2) = \mathbf{E}_i^0 = \mathbf{K}_i^0$, and $\mathbf{E}_i^4$ is bounded by a map $F(\mathbf{E}_i^4) \subset \mathbf{E}_i^2$, whose Hopf invariant[45] is γ_i. Then $\mathbf{K}_0^4$ and $\mathbf{K}_1^4$ have the same natural system, namely the empty system in which R has no generators and a single vacuous relation, and $m_3 = 0$, $m_4 = 1$. It follows from what is known concerning[46] $\pi_3(S^2)$, and S.S., theorem 18, that $\pi_3(\mathbf{K}_i^4)$ is cyclic of order $|\gamma_i|$. Therefore $\mathbf{K}_0^4$ and $\mathbf{K}_1^4$ are not of the same homotopy type if $\gamma_1 \neq \pm\gamma_0$. If $\gamma_0 = 1$, $\gamma_1 = 0$ we may take $\mathbf{K}_0^4$ to be the complex projective plane and $\mathbf{K}_1^4$ to consist of a 4-sphere, $\mathbf{E}_1^4$, and a 2-sphere, $\mathbf{E}_1^2$, with a single common point $\mathbf{E}_1^0$. In this case their homotopy types are also distinguished by the fact that the product[47] of a basic 2-dimensional co-cycle, or surface integral, with itself is co-homologous to zero in the case of $\mathbf{K}_1^4$, and not in the case of $\mathbf{K}_0^4$. Can $\mathbf{K}_0^4$ and $\mathbf{K}_1^4$ be distinguished from each other in this way for arbitrary values of γ_0 and $\gamma_1 \neq \pm\gamma_0$?

[45] See H. Hopf, Math. Annalen, 104 (1931), 637–65.

[46] See H. Freudenthal, Compositio Math., 5 (1938), 299–314.

[47] See Hassler Whitney, Proc. Nat. Academy of Sciences, 23 (1937), 285–91, and papers by E. Čzech and J. W. Alexander, which are referred to there.

III$_4$. *Note on the calculation of* $\pi_r(K)$ $(r > 1)$. We shall say that a group $\mathfrak{G}$, with a countable set of elements, has been *calculated* when it is given in terms of generators and relations, which may be infinite in number, but which must be enumerated. That is to say, a process must be described by which all the generators and relations, may be successively constructed.[48] We shall say that $\mathfrak{G}$ has been *calculated effectively* when it has been calculated, and when a finite algorithm has been provided by which one can tell whether or no two products of the generators represent the same element of $\mathfrak{G}$.

We now show how $\pi_r(K)$ may be calculated for any $r \geqq 1$, where K is a simplicial complex, which may be infinite but locally finite. The group $\pi_1(K)$ can be calculated in various ways, as in §3 for example, and we therefore take $r > 1$. Let $a_1, a_2, \cdots$ be a given, enumerable sequence of vertices and let us define the *weight* of a given symbolic r-simplex $A^r = a_{i_0} \cdots a_{i_r}$ $(i = 1, 2, \cdots)$ as $w(A^r) = i_0 + \cdots + i_r$. We order all the r-simplexes by placing A_1^r before A_2^r if $w(A_1^r) < w(A_2^r)$ and ordering the r-simplexes with a given weight lexicographically. Similarly, we order the purely r-dimensional complexes, $P^r = A_{i_0}^r + \cdots + A_{i_q}^r$, in terms of the r-simplexes. By a *special* r-element we shall mean one which is derived from an r-simplex by reiterating the two operations of stellar sub-division, and addition $E_1^r + E_2^r$, where E_1^r and E_2^r are special r-elements which meet in an $(r-1)$-simplex on the boundary of both. One can find out, by a finite process of trial and error, whether or no a given complex P^r is a special r-element. Therefore the special r-elements are enumerated in the form of a sub-sequence of the purely r-dimensional complexes. Assuming, as we shall do, that the vertices of K are in a given order, one can enumerate the simplicial maps $f(E^r) \subset K$, where E^r is a given r-element, and hence the aggregate of simplicial maps $f_{\lambda i} = f_{\lambda i}(E_i^r) \subset K$ $(\lambda = 1, \cdots, n_i\,;\, i = 1, 2, \cdots)$, such that $f_{\lambda i}(E_i^r)$ is a given vertex $p_0 \in K$, where $E_1^r, E_2^r, \cdots$ are the special r-elements. Any element of $\pi_r(K)$ is represented by such a map, and indeed by one in which E_i^r is a stellar sub-division of an r-simplex. Therefore the set of maps $f_{\lambda i}$ may be taken as the generators of a symbolic group,[49] which, with suitable relations, represents $\pi_r(K)$. For relations we take, first all relations of the form $f_{\lambda i} f_{\mu j} = f_{\rho k} = \cdots = f_{\sigma l}$, which are defined by adding the special r-elements E_i^r and E_j^r, in all possible ways, to form the special r-elements $E_k^r, \cdots, E_l^r$. By means of these relations, which we shall call the relations R, any product of the generators can be transformed into a single generator. Therefore, in addition to the relations R, we only need a set by means of which any representative map, $f_{\rho i}$, of the element $1 \in \pi_r(K)$, can be transformed formally into the constant map f_{1i}, say, together with the relations $f_{1i} = 1$ $(i = 1, 2, \cdots)$. Let A be any internal simplex in E_i^r, let $B = f_\lambda(A)$, for a given value of λ, and let bB be a simplex in K such that, if $a_j A \subset E_i^r$, then $f_\lambda(a_j A) \subset bB$ (the vertex b may belong to B, in which case $bB = B$). Let σ be the elementary

[48] E.g. a process for calculating the n^{th} prime number.

[49] Cf. III$_5$ below. Here we write $\pi_r(K)$ with multiplication, even though $r > 1$.

subdivision (A, a_k), where a_k is the first of the vertices $a_1, a_2, \cdots$ which is not in E_i^r, let $E_j^r = \sigma E_i^r$, and let $f_\mu(E_j^r) \subset K$ be the (simplicial) map which is given by $f_\mu(a_l) = f_\lambda(a_l), f_\mu(a_k) = b$, where a_l is any vertex in E_i^r. Then E_j^r is a special r-element, and $f_{\mu j}$ is one of the maps defined above. Let R' be the set of all relations $f_{1i} = 1, f_{\lambda i} = f_{\mu j}$, where $f_{\lambda i}$ and $f_{\mu j}$ are related in this way. Then, using S.S., Lemma 2 and Theorem 6, it is not difficult to prove that the symbolic group defined by the generators $f_{\lambda i}$ $(\lambda, i = 1, 2, \cdots)$ and the relations R, R′ represents $\pi_r(K)$. The relations R, R′ can obviously be enumerated, whence $\pi_r(K)$ can be calculated for any value of $r = 1, 2, \cdots$. In general it cannot be calculated effectively, even when $r = 1$ and K is finite. Thus if K is infinite $\pi_r(K)$ is on the same logical footing whether $r = 1$ or $r > 1$. If K is finite there is the distinction that $\pi_1(K)$, unlike $\pi_r(K)$ $(r > 1)$, can always be represented by a finite system of generators and relations.

III$_5$. By a symbolic group we mean a group which is originally defined in terms of given generators and relations. An element in such a group is a class[50] of products of the generators and their inverses, and is to be distinguished from an individual product. However we shall use the same symbol to denote a product of generators and the element which it represents. If W_1 and W_2 are two products, then $W_1 = W_2$ is to be interpreted as an equation between elements of a group, unless the contrary is obviously implied by the context, as it is in (4.1), for example.

V. This note contains a simpler example than the one in §5 above, showing that two complexes of the same homotopy type need not have the same nucleus. Let $\mathbf{K}^p$ $(p = 0, 1, 2)$ consist of a p-cell $\mathbf{E}^p$, such that $F(\mathbf{E}^2) = 5\mathbf{E}^1$. Thus the single generator x, and the relation $x^5 = 1$, form a natural system R, for $\pi_1(\mathbf{K}^2)$, and $\tilde{\mathbf{K}}^2$ consists of five 2-cells $x^m\tilde{\mathbf{E}}^2$ $(m = 0, \cdots, 4)$, each bounded by $\sigma\tilde{\mathbf{E}}^1$ $(\sigma = 1 + x + \cdots + x^4)$, where $\tilde{\mathbf{E}}^1 \subset \tilde{\mathbf{K}}^1$ covers $\mathbf{E}^1 \subset \mathbf{K}^1$. A chain $\lambda\tilde{\mathbf{E}}^2$ $(\lambda \in \mathfrak{R})$ is a cycle if, and only if $\lambda\sigma = 0$, and by Hurewicz's theorem any 2-cycle in $\tilde{\mathbf{K}}^2$ is realized by some simplicial map $f(S^2) \subset \tilde{\mathbf{K}}^2$. Therefore $\{\| (x-1)^2 \|, R\}$ is a natural system for a complex $\mathbf{K}^3 = \mathbf{K}^2 + \mathbf{E}^3$, such that $\partial\tilde{\mathbf{E}}^3 = (x-1)^2\tilde{\mathbf{E}}^2$. It may be verified that $x^5 = 1$ implies $(x^3 + x^2 - 1)(x^4 + x - 1) = 1$, whence $\alpha = x^3 + x^2 - 1$ is a regular element of $\mathfrak{R}$. Therefore the system $\{\| \alpha(x-1)^2 \|, R\}$ is L^*-equivalent to $\{\| (x-1)^2 \|, R\}$, and by theorem 9 there is a complex $\mathbf{K}_1^3$, of the same homotopy type as $\mathbf{K}^3$, for which $\{\| \alpha(x-1)^2 \|, R\}$ is a natural system. Let us assume that $\mathbf{K}^3$ and $\mathbf{K}_1^3$ have the same nucleus. Then it follows from theorems 8 and 6 that the sets of incidence matrices

$$\| (x^l - 1)^2 \|, \quad \| \sigma \|, \quad \| x^l - 1 \|$$
$$\| \alpha(x-1)^2 \|, \quad \| \sigma \|, \quad \| x - 1 \|$$

[50] Cf. K. Reidemeister, Einführung in die kombinatorische Topologie, Brunswick (1932), chap. 2.

are L-equivalent for some $l = 1, \cdots, 4$, and from lemma 5 that

$$\alpha(x-1)^3 = \pm x^p(x^l-1)^3,$$

for some value of p. Since

$$x^p(x^l-1)^3 = -x^{p+3l}(x^{-l}-1)^3,$$

we may assume that $l = 1$ or 2. As explained in §5 above, $\lambda(x-1)^3 = 0$ implies $\lambda \equiv 0 \pmod{\sigma}$. Clearly $\alpha \not\equiv \pm x^p \pmod{\sigma}$, whence $l = 2$. Therefore $\alpha \equiv \pm x^p(x+1)^3 \pmod{\sigma}$, whence, writing $x = 1$, we have $1 \equiv \pm 8 \pmod{5}$, which is absurd. Therefore $\mathbf{K}^3$ and $\mathbf{K}_1^3$ do not have the same nucleus.

BALLIOL COLLEGE,
OXFORD.

R.R.
Eglestim

YALE UNIVERSITY
FEb 1 1941
LIBRARY

ANNALS OF MATHEMATICS

(FOUNDED BY ORMOND STONE)

EDITED BY

SOLOMON LEFSCHETZ JOHN VON NEUMANN

FRÉDÉRIC BOHNENBLUST

WITH THE COÖPERATION OF THE

DEPARTMENT OF MATHEMATICS OF PRINCETON UNIVERSITY

AND

THE SCHOOL OF MATHEMATICS OF THE INSTITUTE
FOR ADVANCED STUDY

AND

A. A. ALBERT	T. H. HILDEBRANDT[2]	M. H. STONE
EMIL ARTIN	NATHAN JACOBSON	ANNA PELL-WHEELER[1]
GARRETT BIRKHOFF	SAUNDERS MAC LANE[1]	G. T. WHYBURN[2]
M. R. HESTENES[1]	N. E. STEENROD	OSCAR ZARISKI

([1] Representing the MATHEMATICAL ASSOCIATION OF AMERICA)
([2] Representing the AMERICAN MATHEMATICAL SOCIETY)

SECOND SERIES, VOL. 42, NO. 1
JANUARY, 1941

PUBLISHED QUARTERLY
AT
MOUNT ROYAL AND GUILFORD AVENUES
BALTIMORE, MD.
BY THE
PRINCETON UNIVERSITY PRESS
PRINCETON, N. J.

COPYRIGHT, 1940, BY PRINCETON UNIVERSITY PRESS

ANNALS OF MATHEMATICS

Subscriptions: The subscription price is $5.00 a volume (four numbers); the price of a single copy is $1.50. An arrangement has been made with the MATHEMATICAL ASSOCIATION OF AMERICA whereby an individual member of the ASSOCIATION may subscribe to the ANNALS OF MATHEMATICS at one-half the regular price. Subscription orders, orders for back numbers, notices of change of address, and other business correspondence should be sent to the Princeton University Press, Princeton, N. J.

Editorial Correspondence: Editorial correspondence and manuscripts should be addressed to the ANNALS OF MATHEMATICS, Fine Hall, Princeton, N. J., U. S. A.

Preparation of Manuscripts: Manuscripts submitted to the ANNALS OF MATHEMATICS must be typewritten, double spaced, with generous margins. Figures and diagrams should be drawn in India ink on separate sheets of paper and of twice the size they are to be printed. The lettering is to be omitted from these drawings, but a sketch with complete lettering should be appended.

Footnotes should be numbered consecutively and typed, double spaced, on a sheet by themselves at the end of the manuscript—they will be inserted at the designated places in the text after galley proofs have been corrected. Footnotes should be reduced to a minimum and, whenever possible, replaced by references to a bibliography placed at the end of the paper. Formulae are undesirable in footnotes.

Authors are requested to keep in mind the typographical difficulties involved in setting complicated mathematical formulae. These mechanical difficulties may frequently be eliminated or reduced by the use of alternative notations, *e. g.* "exp" may be used to avoid a complicated exponent, and fractional exponents may be used to replace the troublesome radical sign. The use of unusual symbols is to be avoided. Symbols set one on top of another present a typographical difficulty which usually can be easily avoided by a stepwise arrangement, *e. g.* replace $\sum_{1}^{n}$ by $\sum_{1}^{n}$, etc.

Proofs: Authors ordinarily receive only galley proofs. Stylistic changes in proofs must be avoided; authors will be charged for excessive corrections.

Reprints: Authors receive gratis, postage prepaid, one hundred reprints of each article; joint authors each receive fifty reprints.

Entered as second-class matter at the Post Office at Baltimore, Maryland, under the Act of March 3, 1879.

Made in the United States of America

ANNALS OF MATHEMATICS STUDIES

A new series of lithoprinted monographs (6 x 9) dealing with topics of current mathematical interest and selling at moderate prices.

1. **Algebraic Theory of Numbers.**
   By HERMANN WEYL — 227 *pages*, $2.35
2. **Convergence and Uniformity in Topology.**
   By JOHN W. TUKEY — 95 *pages*, $1.50
3. **The Consistency of the Continuum Hypothesis.**
   By KURT GÖDEL — 68 *pages*, $1.25
4. **An Introduction to Linear Transformations in Hilbert Space.**
   By F. J. MURRAY — (*In preparation*)

The prices above are subject to an educational discount of fifteen per cent.

PRINCETON UNIVERSITY PRESS
Princeton, New Jersey

PRINCETON MATHEMATICAL SERIES

Edited by
MARSTON MORSE H. P. ROBERTSON A. W. TUCKER

1. **The Classical Groups**, their invariants and representations.
   By HERMANN WEYL — 314 *pages*, $4.00
2. **Topological Groups.**
   By L. PONTRJAGIN (translated by E. Lehmer) — 308 *pages*, $4.00
3. **An Introduction to Differential Geometry,** with use of the tensor calculus.
   By LUTHER PFAHLER EISENHART — 300 *pages*, $3.50

IN PREPARATION

4. **Dimension Theory.**
   By WITOLD HUREWICZ & HENRY WALLMAN, *Tentative Price* $2.50
5. **Analytical Foundations of Celestial Mechanics.**
   By AUREL WINTNER

The prices above are subject to an educational discount of fifteen per cent.

PRINCETON UNIVERSITY PRESS
Princeton, New Jersey

Copies of the following memoirs can be obtained by addressing *for the United States, Canada, and South America*, The Nordemann Publishing Company, Inc., 215 Fourth Ave., New York, N. Y.

An introduction to the theory of elliptic functions. By GOSTA MITTAG-LEFFLER. Authorized translation by EINAR HILLE. 81 pages. Price 90 cents.

On various conceptions of correlation. By F. M. WEIDA. 37 pages. Price 75 cents.

Contact transformations. By L. P. EISENHART. 40 pages. Price 80 cents.

A survey of the theory of small samples. By P. R. RIDER. 52 pages. Price $1.00.

Tauberian theorems. By NORBERT WIENER. 100 pages. Price $1.00.

Untersuchungen über Gestalt und Lage abgeschlossener Mengen beliebiger Dimension. By PAUL ALEXANDROFF. 86 pages. Price $1.00.

Plateau's problem and Dirichlet's principle. By RICHARD COURANT. 46 pages. Price 90 cents.

Minimal surfaces of higher topological structure. By J. DOUGLAS. 94 pages. Price $1.00.

Sur quelques notions fondamentales dans la théorie générale des operations linéaires. By FRÉDÉRIC RIESZ. 33 pages. Price 70 cents.

La théorie du corps de classes. By C. CHEVALLEY. 25 pages. Price 50 cents.

now appearing monthly

Mathematical Reviews

(an international journal to abstract and review current mathematical literature)

Sponsored By

American Mathematical Society
Mathematical Association of America
Academia Nacional de Ciencias Exactas Fisicas y Naturales de Lima
and others

subscriptions

send subscriptions to
MATHEMATICAL REVIEWS
Brown University, Providence, R. I., U. S. A.

$6.50 per year to members of sponsoring organizations
$13.00 per year to others

TABLE OF CONTENTS

On Frobeniusean Algebras. II. By TADASI NAKAYAMA 1
Über die Topologie der Gruppen-Mannigfaltigkeiten und Ihre Verallgemeinerungen. By HEINZ HOPF 22
On the Cartan Invariants of Groups of Finite Order. By RICHARD BRAUER 53
Minimal Surfaces Not of Minimum Type by a New Mode of Approximation. By MARSTON MORSE and C. TOMPKINS 62
On a Lemma of McShane. By TIBOR RADÓ 73
On Finsler and Cartan Geometries. III. Two-Dimensional Finsler Spaces with Rectilinear Extremals. By L. BERWALD 84
On Unitary Representations of the Group of De Sitter Space. By L. H. THOMAS 113
Der Dreierstoss. Von CARL LUDWIG SIEGEL 127
Intuitive Probabilities and Sequences. By B. O. KOOPMAN 169
Operator-Theoretical Treatment of Markoff's Process and Mean Ergodic Theorem. By KÔSAKU YOSIDA and SHIZUO KAKUTANI 188
On Types of "Weak" Convergence in Linear Normed Spaces. By INGO MADDAUS, JR. 229
Characterization of Complex Couple Spaces. By ARISTOTLE D. MICHAL and MAX WYMAN 247
Sequences Defined by Non-Linear Algebraic Difference Equations. By OTIS E. LANCASTER 251
Some General Approximation Theorems. By LAWRENCE M. GRAVES 281
Tauberian Conditions. By RALPH PALMER AGNEW 293
On Divergence Properties of the Lagrange Interpolation Parabolas. By P. ERDÖS 309
Additive Set Functions and Vector Lattices. By S. BOCHNER and R. S. PHILLIPS 316
Free Lattices. By PHILIP M. WHITMAN 325
Corrections to Our Paper on the Existence of Minimal Surfaces of General Critical Types. By MARSTON MORSE and C. TOMPKINS 331

The Carus Mathematical Monographs

Thus Far Published

No. 1. *Calculus of Variations*, by PROFESSOR G. A. BLISS. (First Impression, 1925; Second Impression, 1927; Third Impression, 1935.)

No. 2. *Analytic Functions of a Complex Variable*, by PROFESSOR D. R. CURTISS. (First Impression, 1926; Second Impression, 1930.)

No. 3. *Mathematical Statistics*, by PROFESSOR H. L. RIETZ. (First Impression, 1927; Second Impression, 1929; Third Impression, 1936.)

No. 4. *Projective Geometry*, by PROFESSOR J. W. YOUNG. (First Impression, 1930; Second Impression, 1938.)

No. 5. *History of Mathematics in America before 1900*, by PROFESSORS DAVID EUGENE SMITH and JEKUTHIEL GINSBURG. (First Impression, 1934).

Price $1.25 per copy to individual and institutional members of the MATHEMATICAL ASSOCIATION OF AMERICA when ordered through Secretary W. D. Cairns, 97 Elm St., Oberlin, Ohio, (one copy to each member).

Price to non-members, $2.00 per copy, to be ordered from the Open Court Publishing Co., La Salle, Ill.

R. R.
Q

YALE UNIVERSITY
MAY 5 1941
LIBRARY

ANNALS OF MATHEMATICS

(FOUNDED BY ORMOND STONE)

EDITED BY

SOLOMON LEFSCHETZ JOHN VON NEUMANN

FRÉDÉRIC BOHNENBLUST

WITH THE COÖPERATION OF THE

DEPARTMENT OF MATHEMATICS OF PRINCETON UNIVERSITY

AND

THE SCHOOL OF MATHEMATICS OF THE INSTITUTE
FOR ADVANCED STUDY

AND

A. A. ALBERT	T. H. HILDEBRANDT[2]	M. H. STONE
EMIL ARTIN	NATHAN JACOBSON	ANNA PELL-WHEELER[1]
GARRETT BIRKHOFF	SAUNDERS MAC LANE[2]	G. T. WHYBURN[2]
M. R. HESTENES[1]	N. E. STEENROD	OSCAR ZARISKI

([1] Representing the MATHEMATICAL ASSOCIATION OF AMERICA)
([2] Representing the AMERICAN MATHEMATICAL SOCIETY)

SECOND SERIES, VOL. 42, NO. 2

APRIL, 1941

PUBLISHED QUARTERLY
AT
MOUNT ROYAL AND GUILFORD AVENUES
BALTIMORE, MD.
BY THE
PRINCETON UNIVERSITY PRESS
PRINCETON, N. J.

COPYRIGHT, 1941, BY PRINCETON UNIVERSITY PRESS

Subscriptions: The subscription price is $5.00 a volume (four numbers); the price of a single copy is $1.50. An arrangement has been made with the MATHEMATICAL ASSOCIATION OF AMERICA whereby an individual member of the ASSOCIATION may subscribe to the ANNALS OF MATHEMATICS at one-half the regular price. Subscription orders, orders for back numbers, notices of change of address, and other business correspondence should be sent to the Princeton University Press, Princeton, N. J.

Editorial Correspondence: Editorial correspondence and manuscripts should be addressed to the ANNALS OF MATHEMATICS, Fine Hall, Princeton, N. J., U. S. A.

Preparation of Manuscripts: Manuscripts submitted to the ANNALS OF MATHEMATICS must be typewritten, double spaced, with generous margins. Figures and diagrams should be drawn in India ink on separate sheets of paper and of twice the size they are to be printed. The lettering is to be omitted from these drawings, but a sketch with complete lettering should be appended.

Footnotes should be numbered consecutively and typed, double spaced, on a sheet by themselves at the end of the manuscript—they will be inserted at the designated places in the text after galley proofs have been corrected. Footnotes should be reduced to a minimum and, whenever possible, replaced by references to a bibliography placed at the end of the paper. Formulae are undesirable in footnotes.

Authors are requested to keep in mind the typographical difficulties involved in setting complicated mathematical formulae. These mechanical difficulties may frequently be eliminated or reduced by the use of alternative notations, *e. g.* "exp" may be used to avoid a complicated exponent, and fractional exponents may be used to replace the troublesome radical sign. The use of unusual symbols is to be avoided. Symbols set one on top of another present a typographical difficulty which usually can be easily avoided by a stepwise arrangement, *e. g.* replace $\sum_1^n$ by $\sum_1^n$, etc.

Proofs: Authors ordinarily receive only galley proofs. Stylistic changes in proofs must be avoided; authors will be charged for excessive corrections.

Reprints: Authors receive gratis, postage prepaid, one hundred reprints of each article; joint authors each receive fifty reprints.

Entered as second-class matter at the Post Office at Baltimore, Maryland, under the Act of March 3, 1879.

Made in the United States of America

YALE UNIVERSITY
AUG 5 1941
LIBRARY

ANNALS OF MATHEMATICS

(FOUNDED BY ORMOND STONE)

EDITED BY

SOLOMON LEFSCHETZ JOHN VON NEUMANN

FRÉDÉRIC BOHNENBLUST

WITH THE COÖPERATION OF THE

DEPARTMENT OF MATHEMATICS OF PRINCETON UNIVERSITY

AND

THE SCHOOL OF MATHEMATICS OF THE INSTITUTE FOR ADVANCED STUDY

AND

A. A. ALBERT	T. H. HILDEBRANDT[2]	M. H. STONE
EMIL ARTIN	NATHAN JACOBSON	ANNA PELL-WHEELER[1]
GARRETT BIRKHOFF	SAUNDERS MAC LANE[2]	G. T. WHYBURN[2]
M. R. HESTENES[1]	N. E. STEENROD	OSCAR ZARISKI

([1] Representing the MATHEMATICAL ASSOCIATION OF AMERICA)
([2] Representing the AMERICAN MATHEMATICAL SOCIETY)

SECOND SERIES, VOL. 42, NO. 3
JULY, 1941

PUBLISHED QUARTERLY
AT
MOUNT ROYAL AND GUILFORD AVENUES
BALTIMORE, MD.
BY THE
PRINCETON UNIVERSITY PRESS
PRINCETON, N. J.

COPYRIGHT, 1941, BY PRINCETON UNIVERSITY PRESS

ANNALS OF MATHEMATICS

Subscriptions: The subscription price is $5.00 a volume (four numbers); the price of a single copy is $1.50. An arrangement has been made with the MATHEMATICAL ASSOCIATION OF AMERICA whereby an individual member of the ASSOCIATION may subscribe to the ANNALS OF MATHEMATICS at one-half the regular price. Subscription orders, orders for back numbers, notices of change of address, and other business correspondence should be sent to the Princeton University Press, Princeton, N. J.

Editorial Correspondence: Editorial correspondence and manuscripts should be addressed to the ANNALS OF MATHEMATICS, Fine Hall, Princeton, N. J., U. S. A.

Preparation of Manuscripts: Manuscripts submitted to the ANNALS OF MATHEMATICS must be typewritten, double spaced, with generous margins. Figures and diagrams should be drawn in India ink on separate sheets of paper and of twice the size they are to be printed. The lettering is to be omitted from these drawings, but a sketch with complete lettering should be appended.

Footnotes should be numbered consecutively and typed, double spaced, on a sheet by themselves at the end of the manuscript—they will be inserted at the designated places in the text after galley proofs have been corrected. Footnotes should be reduced to a minimum and, whenever possible, replaced by references to a bibliography placed at the end of the paper. Formulae are undesirable in footnotes.

Authors are requested to keep in mind the typographical difficulties involved in setting complicated mathematical formulae. These mechanical difficulties may frequently be eliminated or reduced by the use of alternative notations, *e. g.* "exp" may be used to avoid a complicated exponent, and fractional exponents may be used to replace the troublesome radical sign. The use of unusual symbols is to be avoided. Symbols set one on top of another present a typographical difficulty which usually can be easily avoided by a stepwise arrangement, *e. g.* replace $\sum_{1}^{n}$ by $\sum_{1}^{n}$, etc.

Proofs: Authors ordinarily receive only galley proofs. Stylistic changes in proofs must be avoided; authors will be charged for excessive corrections.

Reprints: Authors receive gratis, postage prepaid, one hundred reprints of each article; joint authors each receive fifty reprints.

Entered as second-class matter at the Post Office at Baltimore, Maryland, under the Act of March 3, 1879.

Made in the United States of America

YALE UNIVERSITY
NOV 3 1941
LIBRARY

ANNALS OF MATHEMATICS

(FOUNDED BY ORMOND STONE)

EDITED BY

SOLOMON LEFSCHETZ JOHN VON NEUMANN

FRÉDÉRIC BOHNENBLUST

WITH THE COÖPERATION OF THE

DEPARTMENT OF MATHEMATICS OF PRINCETON UNIVERSITY

AND

THE SCHOOL OF MATHEMATICS OF THE INSTITUTE FOR ADVANCED STUDY

AND

A. A. ALBERT	T. H. HILDEBRANDT[2]	M. H. STONE
EMIL ARTIN	NATHAN JACOBSON	ANNA PELL-WHEELER[1]
GARRETT BIRKHOFF	SAUNDERS MAC LANE[2]	G. T. WHYBURN[2]
M. R. HESTENES[1]	N. E. STEENROD	OSCAR ZARISKI

([1] Representing the MATHEMATICAL ASSOCIATION OF AMERICA)
([2] Representing the AMERICAN MATHEMATICAL SOCIETY)

SECOND SERIES, VOL. 42, NO. 4
OCTOBER, 1941

PUBLISHED QUARTERLY
AT
MOUNT ROYAL AND GUILFORD AVENUES
BALTIMORE, MD.
BY THE
PRINCETON UNIVERSITY PRESS
PRINCETON, N. J.

COPYRIGHT, 1941, BY PRINCETON UNIVERSITY PRESS

Subscriptions: The subscription price is $5.00 a volume (four numbers); the price of a single copy is $1.50. An arrangement has been made with the MATHEMATICAL ASSOCIATION OF AMERICA whereby an individual member of the ASSOCIATION may subscribe to the ANNALS OF MATHEMATICS at one-half the regular price. Subscription orders, orders for back numbers, notices of change of address, and other business correspondence should be sent to the Princeton University Press, Princeton, N. J.

Editorial Correspondence: Editorial correspondence and manuscripts should be addressed to the ANNALS OF MATHEMATICS, Fine Hall, Princeton, N. J., U. S. A.

Preparation of Manuscripts: Manuscripts submitted to the ANNALS OF MATHEMATICS must be typewritten, double spaced, with generous margins. Figures and diagrams should be drawn in India ink on separate sheets of paper and of twice the size they are to be printed. The lettering is to be omitted from these drawings, but a sketch with complete lettering should be appended.

Footnotes should be numbered consecutively and typed, double spaced, on a sheet by themselves at the end of the manuscript—they will be inserted at the designated places in the text after galley proofs have been corrected. Footnotes should be reduced to a minimum and, whenever possible, replaced by references to a bibliography placed at the end of the paper. Formulae are undesirable in footnotes.

Authors are requested to keep in mind the typographical difficulties involved in setting complicated mathematical formulae. These mechanical difficulties may frequently be eliminated or reduced by the use of alternative notations, *e. g.* "exp" may be used to avoid a complicated exponent, and fractional exponents may be used to replace the troublesome radical sign. The use of unusual symbols is to be avoided. Symbols set one on top of another present a typographical difficulty which usually can be easily avoided by a stepwise arrangement, *e. g.* replace $\sum_{1}^{n}$ by $\sum_{1}^{n}$, etc.

Proofs: Authors ordinarily receive only galley proofs. Stylistic changes in proofs must be avoided; authors will be charged for excessive corrections.

Reprints: Authors receive gratis, postage prepaid, one hundred reprints of each article; joint authors each receive fifty reprints.

Entered as second-class matter at the Post Office at Baltimore, Maryland, under the Act of March 3, 1879.

Made in the United States of America